MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

EDITED BY

PROF. GILBERT RYLE

WITH THE CO-OPERATION OF PROF. SIR F. C. BARTLETT AND PROF. C. D. BROAD

CONTENTS

Price Seven Shillings and Sixpence   *All Rights Reserved*

Yearly subscribers will receive MIND *post free* from the Publishers on payment (in advance) of Twenty-five Shillings.

Printed in Great Britain

THE
Philosophical
Review

ed
e a
cal
ent

to
an
net

ng
ies
net

ns
of
sh
es
net

—

ce
y
er
an
ce
el
ll

V

P

T<small>HE</small>
pu
are
up
ess
tai
As
six
sel
and
top
nat
of
obl
pro

I
one
kin
gav
nat
tic
(ii)
to s
Moc
that

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

————— ∞ε —————

I.—G. E. MOORE'S LATEST PUBLISHED VIEWS ON ETHICS

By C. D. Broad

THE first six essays in the book *The Philosophy of G. E. Moore*, published in 1942 as Vol. IV in *The Library of Living Philosophers*, are devoted to Moore's ethical theories ; and Moore's comments upon them occupy the first ninety-three pages of his terminal essay. I suppose that this part of the terminal essay must contain Moore's latest published pronouncements on ethical problems. As such, it is of considerable interest and importance. Of the six ethical essays and Moore's comments on them I propose to select three for discussion here, *viz.* those of Frankena, Stevenson, and myself. Between them they cover the following four main topics, *viz.* (I) The distinction between ' natural ' and ' non-natural ' characteristics, (II) The ' autobiographical ' analysis of moral indicatives, (III) The interconnections of value and obligation, and (IV) Ethical egoism and ethical neutralism. I propose to treat each of these topics in turn.

(I) *The distinction between ' natural ' and ' non-natural ' characteristics*

It is a well known doctrine of Moore's that the word ' good ', in one important sense of it, stands for a characteristic of a peculiar kind which he terms ' non-natural '. In *Principia Ethica* he gave certain criteria for distinguishing ' natural ' and ' non-natural ' characteristics. The two marks of a *natural* characteristic were said to be (i) that it ' can exist in time all by itself ', and (ii) that it is a ' part ' of anything that it characterizes. I tried to show in my essay that these criteria are utterly unsatisfactory. Moore accepted that criticism ; and so we may henceforth regard that part of his doctrine as withdrawn.

In my essay I suggested that Moore was almost certainly intending to deal with the same distinction (though he does not use the words 'natural' and 'non-natural') in the paper entitled "The Conception of Intrinsic Value" in his *Philosophical Studies* (1922). I understood his doctrine there to be as follows. (1) The characteristics of a thing may be divided into (a) those that *do*, and (b) those that do *not*, 'depend solely on its intrinsic nature'. (2) Those characteristics which do depend solely on the intrinsic nature of that which they characterize may be subdivided into (α) those which *are*, and (β) those which are *not* 'intrinsic'. (3) The *non-natural* characteristics of a thing are the members of the sub-class (a, β), *i.e.* those which *are* dependent solely on its intrinsic nature but are *not* intrinsic. The *natural* characteristics of a thing are the members of class (b) and the members of sub-class (a, α), *i.e.* they are those characteristics of it which *either* do not depend solely on its intrinsic nature *or* which depend solely on its intrinsic nature and are also intrinsic.

In his terminal essay Moore points out where I was right and where I was wrong in my interpretation of his doctrine in "The Conception of Intrinsic Value." I was right in thinking that he was concerned there with the distinction which he described in *Principia Ethica* by the words 'natural' and 'non-natural'. But I was wrong in thinking that he would admit there to be such a class of characteristics as (a, β), *i.e.*, ones which *do* depend solely on the intrinsic nature of that which they characterize and yet are *not* intrinsic. Moore says that he held that *all* characteristics which depend solely on the intrinsic nature of that which they characterize *are intrinsic*. And he held that goodness, in the fundamental sense in which he is here concerned with it, is intrinsic.

He thinks that my mistake may have arisen from the very unfortunate terminology which he used in "The Conception of Intrinsic Value." He admits that he there used the term 'intrinsic property' in such a way that there would be no inconsistency between the following three statements, (i) 'P is intrinsic', (ii) 'P is a property', and (iii) 'P is not an intrinsic property'. For, he says, his doctrine was that goodness (in the sense in question) *is* intrinsic and *is* a property and yet *is not* an intrinsic property of a good thing.

In view of this, I think that my misunderstanding was not only excusable but also fortunate, for it gave Moore an opportunity to remove what must have been a constant source of confusion even to wary readers. Henceforth, he says in the terminal essay, he will drop this terminology. In future, if I understand him aright,

he would call *all* those properties and *only* those properties of a thing, which depend solely on its intrinsic nature, ' intrinsic properties ' of it. He would then sub-divide the intrinsic properties of a thing into ' natural ' and ' non-natural '. And he would hold that goodness (in the sense in question) is a non-natural intrinsic property of a good thing. It will be noted that ' being an intrinsic property of a thing ' is defined in terms of the notion of ' depending solely on the intrinsic nature of a thing '. The latter notion is elaborately expounded in " The Conception of Intrinsic Value ". I did not criticize it in my essay, and Moore takes it for granted in his terminal essay ; so I shall not discuss it here.

The verbal confusion is now removed, but we are left with the substantial question : What is Moore's criterion for distinguishing between those intrinsic properties of a thing which are *natural* and those which are *non-natural*? In " The Conception of Intrinsic Value " Moore gave two criteria, and the first of these may be subdivided into two complementary parts. In the amended terminology they may be stated as follows. (1.1) A complete enumeration of the *natural* intrinsic characteristics of a thing would be a *complete description* of that thing. (1.2) An enumeration which omitted any *natural* intrinsic characteristic of a thing would be an *incomplete description* of that thing. (2) The *natural* intrinsic characteristics of a thing seem to contribute towards describing its intrinsic nature in a way in which predicates of value do not.

In my essay I confined myself to (1.1), and said nothing about (1.2) or about (2). Moore admits in the terminal essay that (1.2) cannot be maintained as it stands. Suppose that P and π are two properties, *e.g.* being red and being coloured, such that anything that had P would, as a necessary consequence, have π. Then a description which included P would not be made incomplete merely by omitting π. And yet π might be a *natural* intrinsic characteristic. So (1.2) would have to be amended to run somewhat as follows : No description of a thing would be complete, if it omitted any *natural* intrinsic characteristic of it which is not conveyed by some one or some combination of its other natural characteristics. (I use the phrase ' P *conveys* Q ' to mean the same as ' If anything had P, it would necessarily follow from that alone that it would have Q '.)

Moore admits in the terminal essay that he did not clearly distinguish criteria (1.1) and (1.2), on the one hand, from criterion (2), on the other. He says that he is now inclined to rely mainly on the following amended form of (2), *viz.* that, in one sense of

' describe ', the mention of *any natural* characteristic of a thing
contributes to some extent to describe that thing ; whilst the
mention of its *non-natural* intrinsic characteristics does not, in
that sense, describe it at all. He admits that this is extremely
vague, unless we can give some more definite information as to
the particular sense of ' describe ' which is here relevant.

I think it is fair to conclude that Moore, at the time when he
wrote this terminal essay, was unable to give any satisfactory
definition of, or criterion for, a ' non-natural characteristic '.
But I think that we can go further. His suggested criterion,
with its admitted vagueness, due to the uncertainty of the rele-
vant sense of ' describe ', is surely grist to the mill of supporters
of what I will call ' non-predicative interpretations of moral
sentences in the indicative '. If, as that theory holds, the word
' good ' is not the name of a characteristic at all, but its use is,
e.g. to express or to evoke certain emotions, then to call a thing
' good ' would *not* contribute in any way to the description of it.
And yet, owing to the likeness of grammatical form between such
sentences as ' That is a pleasant emotion ' and ' That is a morally
good emotion ', *e.g.* there might well seem to be something para-
doxical in saying that the former did, and the latter did not,
contribute towards describing the emotion. So one could under-
stand why those who never questioned that moral sentences in
the indicative assign a predicate to a subject, should sum up the
situation by alleging that the word ' good ' stands for a property
of a peculiar kind, which does not contribute to describe its
subject in the familiar way in which *e.g.* the property denoted by
' pleasant ' does.

There remain two small points which are worth mentioning
before leaving this part of the subject. (1) Moore says that,
in his opinion, there are at least two kinds of intrinsic value, *viz.*
goodness (in the sense in question) and beauty. But he does not
hold, and never has held, that goodness, in that sense, is a deter-
minable in W. E. Johnson's usage of that word. I must say, for
my own part, that I should need a great deal of persuasion before
I would admit that there is even a *prima facie* case for regarding
beauty, in any sense of that word, as a form of *intrinsic* value.

(2) The other point is this. In the course of my essay I used
an argument which presupposes that the pleasantness of a pleasant
experience is dependent solely on its intrinsic nature. I assumed,
e.g. that, in the case of a pleasant sensation, its pleasantness is
always conveyed by some intrinsic pleasant-making sensible
quality of it, such as its sweetness. Now Moore points out that
the relation between the pleasant-making characteristics of an

experience and its pleasantness is almost certainly *not* that of conveyance, but is that of *causal determination*. I fully agree with that contention, and I will proceed to develop it in my own way, in which Moore might not have been willing to follow.

The essential point is that it is perfectly conceivable that two persons, or the same person on different occasions, should have sensations which were exactly alike in all their sensible qualities, and yet that one of them should be a pleasant experience and the other an unpleasant one. It is a very well founded empirical generalization, *e.g.* that the vast majority of human beings, whenever they have a sensation of the ' toothachy ' kind, dislike that sensation for its characteristic sensible qualities. That is why we call toothachy sensations ' unpleasant '. But there is no kind of necessity about that generalization. It is perfectly conceivable that there might be persons who, when they had a sensation of precisely the same kind, always, or on certain special occasions, *liked* that sensation for those very same sensible qualities for which most persons at most times *dislike* such sensations. For any such person, on any such occasion, a toothache would be a *pleasant* experience. I would suggest, then, that the words ' pleasant ' and ' unpleasant ', as applied to experiences, often imply a well-founded empirical generalization, to the effect that the vast majority of people, on the vast majority of occasions when they have an experience of a certain kind, would like it (or dislike it, as the case may be) for its characteristic experiential qualities. But there is also, plainly, a non-statistical sense of the words ' pleasant ' and ' unpleasant '. To call an experience ' pleasant ' (or to call it ' unpleasant '), in this latter sense, means that the particular person, who has it on a particular occasion, then and there *likes* it (or *dislikes* it, as the case may be) for its characteristic experiential qualities. There is no kind of contradiction in saying that a particular experience, which would correctly be called ' pleasant ' (or be called ' unpleasant ') in the *statistical* sense, occurring on a particular occasion in a particular person, might be correctly called ' unpleasant ' (or ' pleasant ', as the case may be) in the *non-statistical* sense.

(II) *The ' autobiographical ' Analysis of Moral Indicatives*

Consider the sentence : ' It was right for Brutus to stab Caesar ', uttered at a certain moment by a person who is really considering what he is saying and is not merely talking like a parrot or giving an example in an essay. What I call the ' autobiographical ' analysis of this sentence is, on its positive side,

that the speaker is intending to state, beside the historical proposition that Brutus stabbed Caesar, the autobiographical proposition that he himself is feeling a certain kind of emotion (*viz.* one of moral approval) in contemplating that historical proposition. On the negative side the theory is that the speaker is not intending to state anything else beside that historical and that autobiographical proposition and anything that may be logically entailed by them.

This must be carefully distinguished from what I have called ' the non-predicative theory of moral indicatives '. That holds that the speaker is stating nothing but the historical proposition; but that he would not have used the moral-indicative form of expression unless he were feeling moral approval towards it himself or had wanted to induce that emotion in his hearers. (The theory can, of course, take other forms, with something else substituted for ' moral approval '.) None of the essayists explicitly defends the non-predicative theory. But Professor Stevenson defends the autobiographical analysis against certain arguments which Moore had used in his paper " The Nature of Moral Philosophy " in *Philosophical Studies*. Moore, in his reply, says that he would be more inclined to accept the non-predicative theory than the autobiographical analysis, if he were to accept either.

Before going further it is worth while to note that the auto-biographical analysis might take two different forms, which I will call ' occurrent ' and ' dispositional '. On the occurrent form of it, a person who says at a certain moment that X is right is saying that he is at that moment feeling moral approval for X. On the dispositional form of the theory, he is saying that he is generally disposed to feel moral approval when he contemplates actions like X. Moore distinguished those two forms of the theory in his paper " The Nature of Moral Philosophy ". But Stevenson considered only the occurrent form, and therefore Moore also confines himself to that in his reply. This seems to me unfortunate, because the dispositional form is much more plausible than the occurrent form.

There is a matter, which seems to me quite simple, about which both Stevenson and Moore make terribly heavy weather. The essential point at issue can be put as follows. Suppose that the occurrent form of the autobiographical analysis were correct. Then A's utterance at t of a token of the type-sentence ' X is right ' would be equivalent in meaning to his uttering a token of the type-sentence ' I am now feeling moral approval of X '. Similar remarks apply, *mutatis mutandis*, to ' X is wrong '. Now

that makes the predicates ' right ' and ' wrong ' to be doubly relational, for it makes them involve a relation to a speaker and to a time. It follows that ' right ' would have a systematically different meaning on every different occasion on which it is predicated, even by the same person, beside having a systematically different meaning corresponding to each different person who predicates it on any occasion. Now the word ' right ' seems *prima facie* not to answer to those conditions. It seems to be used as if it could stand for precisely the same characteristic when predicated by different persons or on different occasions by the same person. Moore's arguments against the occurrent autobiographical analysis in " The Nature of Moral Philosophy " are simply various ways of trying to exhibit strikingly certain aspects of this *prima facie* conflict between the common usage of the word ' right ' and the usage which would seem to be required if the occurrent autobiographical analysis were correct.

One of Moore's arguments was concerned with the possible alteration in a person's emotional attitude towards the same action, if he should contemplate it on successive occasions. Stevenson's criticism of this argument brings out an important point about the use of tenses in such sentences as ' X *is* right ', ' X *was* right ', and ' X *will be* right '. The point may be put as follows. Suppose that A says at *t* ' I now approve of X, but I formerly contemplated it with disapproval '. Obviously his statement may be true. Now Moore has argued that, if the occurrent autobiographical analysis be correct, A's statement would be equivalent to ' X is now right, but was formerly wrong '. And he had pointed out that it is nonsensical to say, of *one and the same* action, that it was right at one time and became wrong later.

Now Stevenson quite justifiably challenges Moore's right to assert that the theory entails the equivalence mentioned above. Stevenson insists that the correct interpretation of the theory is as follows. If a person says ' X *is* right ', he means that he is *now* feeling approval towards X, which is *now* being performed. If he says ' X *was* right ', he means that he is *now* feeling approval towards X, which *has been* performed. The tense in the moral indicative refers only to the date of the action which is said to be right or to be wrong; and the principal tense in the autobiographical equivalent of that indicative is *always the present*.

If we accept this contention of Stevenson's, what really does follow from the autobiographical analysis, together with the fact of the change in A's attitude, is this. A can now correctly and truly say ' X was right '; and he could, at some former time, have said with equal correctness and equal truth ' X is (or was)

wrong '. But at no time could he correctly and truly say ' X was right at one time and is now wrong '. For that would be equivalent to uttering the sentence ' I now approve of X, which happened in the past, and disapprove of X, which is happening now '. This is doubly nonsensical, since it asserts that the speaker had, at the same time, incompatible emotional attitudes towards one and the same particular action, and it implies that one and the same action was done at two different times.

Another argument in Moore's paper on " The Nature of Moral Philosophy " may be put as follows. Suppose that A and B contemplate the same act X at the same time t. A may say ' I approve of X ', and B may say ' I disapprove of X ', and both may be telling the truth. Now, if the analysis under discussion be correct, A's statement is equivalent to his saying, ' X is right', and B's statement is equivalent to his saying ' X is wrong'. Now the two latter statements conflict logically, whilst the two former are logically compatible. Therefore they cannot be equivalent each to each.

The true account of this situation is admirably brought out by Moore in his terminal essay. It is this. If the analysis under discussion be admissible, A can correctly and truly say ' X is right', and B can at the same time correctly and truly say ' X is wrong '. But no-one at any time can correctly and truly say ' X is both right and wrong '. For anyone who did so would, according to the proposed analysis, be saying ' I now approve and disapprove of X '. Now that could not be truly said by A, who approves and does not disapprove of X ; nor by B, who disapproves and does not approve of X ; nor by any third person, since no-one can entertain simultaneously incompatible emotional attitudes towards the same object.

This amendment, however, does nothing to diminish the force of Moore's original argument against the occurrent autobiographical analysis, viz., that, according to it, A and B do not differ in opinion when one of them pronounces an action to be right and the other pronounces the very same action to be wrong. This is recognized by Stevenson, who proceeds to meet it by making two additions to the proposed analysis.

The first is to point out that, although A and B would not differ in opinion, in the sense of holding incompatible beliefs, they would do so in the wider sense of having opposed emotional attitudes toward the same object. The second is to remind us that in such situations each person would generally seek to alter the emotional attitude of the other and make it resemble his own.

Stevenson admits that, even when due weight has been given to these two considerations, the occurrent autobiographical analysis is not wholly satisfactory. Suppose that A asks B ' Is X right? ' A is not as a rule wanting to find out whether *he himself* now approves of X, but whether B or most other people would do so. Or, again, A may disapprove of X and may know that B approves of it, and the motive of his question may be to induce B to change his attitude. Lastly, if A asks *himself* ' Is X right? ', he is certainly not trying to find out whether he now approves of X. The situation probably is that he has conflicting attitudes towards X, in respect of various aspects of it, and that he is seeking to straighten them out.

Moore does not seriously dispute anything that Stevenson here says. He tells us that he has always recognized that difference of ' opinion ' covers opposition of emotional attitude, but that he used not to think it possible that moral conflicts could be merely of that kind. He is now inclined to think that moral disagreement *may* be nothing but opposition of emotional attitude ; but he is also inclined about equally strongly to think that it involves a logical conflict between incompatible beliefs. Stevenson, he says, has given no reasons for his own alternative ; he has merely shown that certain arguments against it are inconclusive. If Moore felt obliged to abandon his own theory, he would not be inclined to stop at the stage of the occurrent autobiographical analysis, but would prefer to accept some form of *non-predicative* theory. Moore says that he is, in fact, now quite strongly disposed to think that, when a person utters the sentence ' X is right ', he is not asserting *anything* that could be true or false, not even the autobiographical proposition that he now approves of X. But Moore says that he also continues to have some inclination to hold his old view. And he cannot say which of these inclinations is the stronger.

(III) *The interconnections of Value and Obligation*

The longest and most complex essay in the ethical part of the book is that of Professor Frankena, and the part of Moore's reply which deals with it is also highly involved. The question at issue is the connection between the fact that a state of affairs would be *intrinsically good* and a person's being under an *obligation* to seek to *bring it into existence*. Moore had made certain statements on this topic in his various ethical writings, and Frankena discusses their truth and their compatibility with

Moore's characteristic doctrines that good is a simple, indefinable, intrinsic, and non-natural characteristic.

The best way to convey an idea of the discussion is to take in turn the points which Frankena enumerates in the summary at the end of his essay, and to consider, in each case, Moore's treatment of them.

Point 1. This divides into two propositions, which I will call (1, *a*) and (1, *b*). The former is the contention that, if good (in Moore's sense) be *simple*, then the statement ' I am morally bound to do Y ' cannot *mean* the same as the statement ' Y will produce more good or less evil than any other act open to me '. The latter is the contention that the same negative consequence follows from the supposition that good (in Moore's sense) is *intrinsic*, in the sense explained by him.

(1, *a*) After a good deal of discussion on alleged obscurities and ambiguities in Frankena's reasoning, Moore proceeds to state formally what he takes to be Frankena's argument on this point. I have very little doubt that this is a correct account of what was present, in a less precise form, in Frankena's mind, and so I shall adopt it. The argument may be stated as follows. The proposition that good is *simple* entails that statements of the form ' X is good ' neither include nor are identical with statements of obligation. That entails that statements of the form ' X is good ' are *not normative*. That in turn entails that statements of the form ' Y will produce the most good or the least evil of all the acts open to me ' are *not normative*. And that entails that statements of the latter form are not identical in meaning with statements of the form ' I ought to do Y '.

Now Moore holds that the fundamental step in this argument is the second, *viz.* that if statements of the form ' X is good ' neither include nor are identical in meaning with statements of obligation it follows that such statements are not normative. The validity or invalidity of this step depends on what Frankena means by ' normative ', and that (Moore alleges) is not made perfectly clear in his essay. But, setting aside minor verbal inconsistencies, it seems fairly plain that what he intends is the following. S is a normative statement about an action, if and only if it follows from the very nature of that statement that that action ought to be done. If we accept this account of ' normative ', we see that the transition in step 2 depends on the tacit assumption that nothing can follow from the very nature of a statement except what is identical with or is a part of what is meant by the latter. There is in fact no doubt that Frankena does assume this premiss, for elsewhere in his essay he makes it

quite explicit that he thinks that the two propositions ' Q *follows from the very nature of* P ' and ' Q is *synthetically, though necessarily,* connected with P ' are mutually exclusive. Now Moore rejects this premiss, and therefore sees no reason to accept step 2 of Frankena's argument.

As this is an important point, I will state all that is to be found in Moore's terminal essay on this topic. In the first place, he gives an example taken from Professor Langford's essay in the same volume. He says that, in his opinion, it *does* follow, from the very nature of the statement ' This is a cube ', that this has twelve edges ; whilst the latter is *not* identical with nor a part of the meaning of the former.   Secondly, in another part of his essay, Moore makes the following general assertions. He says that he uses the phrase ' Q *follows from* P ' to mean that the conjunction ' P & not-Q ' is *self-contradictory*. But he holds that such a conjunction may be self-contradictory *without* ' Q follows from P ' being *analytic*.

If we put all this together, we see that what Moore is maintaining is the following. Even though good be *simple*, the conjunction ' I ought to do Y, and Y will *not* produce as good consequences as some other action open to me ' may be *self-contradictory*, in that sense (whatever it may be) in which the conjunction ' This is a cube, and this has *not* twelve edges ' is self-contradictory. I should agree that this is quite possible, provided that *ought* itself is not simple, but contains *good* in its analysis.   But, if *good* and *ought* were both simple, I cannot for the life of me see how the conjunction in question could be *self-contradictory*, in any generally accepted sense of that phrase.   It might, however, be *self-evidently impossible*, without being self-contradictory in the formal sense, if we admit the possibility of necessary connections and disconnections which are synthetic, but obvious on inspection. I should add, perhaps, that I am extremely doubtful whether the conjunction ' This is a cube, and this has not twelve edges ' *is* self-contradictory. I should suspect that what is so is the conjunction of this conjunction with certain of the axioms of three-dimensional Euclidean geometry. If so, it is not very helpful as an analogy to the ethical propositions under consideration.

Moore remarks that Frankena might reply to his criticisms on step 2 of the argument by saying that he uses the word ' include ' in such a way that ' Q is included in P ' covers *inter alia* ' Q follows necessarily but synthetically from P '. But that would not help Frankena's argument, since it would save step 2 only at the expense of step 1. For, if ' include ' be used in this

extended sense, there is no reason why the simplicity of *good*
should prevent statements of the form ' X is good ' from
' including ' statements of obligation.

(1, *b*) This is the contention that, if good be *intrinsic* in Moore's
sense of that word, then the statement ' I am morally bound to
do Y ' cannot mean what is meant by ' Y will produce more good
or less evil than any other act open to me '.

Moore says that the argument is precisely the same as that in
(1, *a*), with ' intrinsic ' substituted for ' simple '. It therefore
suffers from the same defect, *viz.* that the second step is un-
justified, for the reasons given above. But it suffers from a
further defect. For the first step, which was quite legitimate in
(1, *a*), ceases to be so when ' intrinsic ' is substituted for ' simple '.
From the hypothesis that good is *intrinsic*, in Moore's sense, it
would *not* follow that statements of the form ' X is good ' neither
include nor are identical with statements of obligation.

In order to discuss this, we must remember what Moore does
and what he does not mean by calling a characteristic ' intrinsic '.
To say that P is an intrinsic characteristic of X means that the
possession of P by X depends solely on X's intrinsic nature.
Now, in the first place, it does *not* follow from this definition that
every intrinsic characteristic of X must be a *pure quality*. No
doubt, if goodness were a pure quality, whether intrinsic or not,
it would follow at once that ' X is good ' could not be identical
with or include any statement of obligation. For the latter
would involve *relations* to an actual or possible agent. But Moore
has distinguished between the ' external ' and the ' internal '
relational properties of a thing ; and, whilst no *external* relational
property of a thing could be intrinsic, there is nothing to prevent
its *internal* relational properties from being so.

We may put the matter as follows. We must distinguish
between what we might call ' categorical ' and ' conditional '
relational properties, though Moore does not use those terms. It
would be a *categorical* relational property of a certain bit of
arsenic to be poisoning Mr. Jones at a certain moment. That
property would be external and non-intrinsic ; for that bit of
arsenic would not be having it unless Mr. Jones had existed and
had swallowed it. It is a *conditional* relational property of any
bit of arsenic to be poisonous, *i.e.* to be such that it *would* poison
a man, *if* there were one and *if* he were to swallow it. This
property, though relational, may be internal and intrinsic ; for a
bit of arsenic would have it even though there had never been any
men or though no man had ever swallowed it. Similarly, if good-
ness be an *intrinsic* property of X, the statement ' X is good '

cannot include or be identical in meaning with any such *categorical* statement as ' Z ought to desire X ' or ' Z ought to try to produce X '. But there is nothing to prevent its including or being identical with some *conditional* proposition of the form ' If there were a person who fulfilled such and such conditions, he would be under an obligation to desire X or to try to produce X '. For X could have that property, even if there had never been any persons, or if no person had ever fulfilled the required conditions.

Whilst I admit the validity and the importance of the distinction which Moore draws here, I do think that it is rather misleading to say of even a *conditional* relational property that it ' depends solely on the intrinsic nature of its possessor '. Surely there is an important sense in which the poisonousness of arsenic depends just as much on the intrinsic nature of *a living organism* as on the intrinsic nature of arsenic. In the same sense and to the same degree the property of being such that, if there were a person and he were to fulfil certain conditions, he would be under an obligation to try to produce X, depends just as much on the intrinsic nature of *moral persons* as on that of X. No doubt arsenic would have been poisonous, even if there never had been and never will be any living organisms ; but at least we can say that the very notion of poisonousness involves the notion of organisms and vital processes, and that no amount of reflection on arsenic in isolation could have supplied the latter notions.

Point 2. Frankena's second point really divides into seven interconnected propositions. It may be stated as follows. If value be either (*a*) *simple*, or (*b*) *intrinsic*, then it cannot be either (α) *normative*, or (β) *non-natural*, or (γ) *definable in terms of obligation*. And, that being so, (*c*) there is no reason to think that it is *incapable of being defined in non-ethical terms*.

It is evident that we thus have six hypothetical propositions, which arise by combining in turn each of the two antecedents (*a*) and (*b*) with each of the three consequents (α), (β), and (γ). In addition to these six hypotheticals there is the seventh proposition (*c*), which Frankena states in the form ' In that case there is no reason to regard value as being incapable of definition in non-ethical terms '. We may label the six hypotheticals as (2*a*, α), (2*a*, β), (2*a*, γ), and (2*b*, α), (2*b*, β), (2*b*, γ). The seventh proposition may be labelled (2*c*).

Moore claims to have dealt with (2*a*, α), (2*a*, γ), (2*b*, α), and (2*b*, γ) in his discussion of (1, *a*) and (1, *b*). He has admitted (2*a*, γ), *i.e.* that, if good be simple, it cannot be defined in terms of obligation, since it would be indefinable. He has rejected (2*a*, α),

$(2b, \alpha)$, and $(2b, \gamma)$. It remains, therefore to deal with $(2a, \beta)$, $(2b, \beta)$, and $(2c)$. That we will now proceed to do.

$(2a, \beta)$. This is the proposition that, if good be *simple*, it cannot be *non-natural*. The essence of Frankena's argument is as follows. If good were simple, it would not be normative. If it were not normative, there would be no reason to think it non-natural. Therefore, if it were simple, it would be non-natural.

Now the first step has already been discussed and rejected. And, even if both it and the second step were accepted, the correct conclusion would be only that, if good were simple, there would be *no reason to think* that it is non-natural. There would be no justification for the stronger conclusion that it *would not be* non-natural.

$(2b, \beta)$. This is the proposition that, if good be *intrinsic*, it cannot be *non-natural*. The argument is the same as before, with ' intrinsic ' substituted throughout for ' simple '. The first step of this argument has already been discussed and rejected. And the argument has the same defect as $(2a, \beta)$, *viz.* that of drawing a stronger conclusion than would be justified by its premisses, even if these were acceptable.

Before passing to $(2c)$, it will be worth while to consider for a moment the second premiss, which is common to both the above arguments of Frankena's. This is the proposition that, unless good were *normative*, there would be no reason to think it *non-natural*. The essence of Frankena's contention on this topic is as follows. In his opinion, the main point of the doctrine that intrinsic value is non-natural is that it cannot be reduced to purely psychological, sociological, biological, or metaphysical terms. Now it seems to him that the only feature in moral judgments which can plausibly be held not to be so reducible is their ostensibly *normative* character, *i.e.* ' the fact that they seem to be saying of some agent that he *ought* to do something '. He concludes that, unless intrinsic value ' in itself possesses a normative character or obligatoriness ', there is no reason to think that it is not essentially reducible to the terms enumerated above.

$(2c)$. This proposition, which comes immediately after the six hypotheticals which we have now discussed, is stated in the very obscure sentence : ' In that case there is no reason to think that good is not definable in non-ethical terms.' We naturally ask : ' In *what* case? ' In the context Frankena might mean *either* that if his six hypotheticals were true there would be no reason to think that good is not definable in non-ethical terms, *or* that if their three consequents were true there would be no reason to think this. Moore does not consider the first of these alternatives, but confines his attention to the second. This is the

proposition that, if good be neither normative nor non-natural nor definable in terms of obligation, then there is no reason to think it is not definable in non-ethical terms.

The phrase ' not definable in non-ethical terms ' needs a certain amount of unpacking. It will be best to start from the beginning. Good might be either (1) indefinable, or (2) definable. If indefinable, it might be either (1.1) identical with some admittedly non-ethical simple notion, *e.g.* pleasant, or (1.2) not identical with any admittedly non-ethical notion. If definable, it might be either (2.1) definable in wholly non-ethical terms, or (2.2) definable only in terms which are wholly or partly ethical. We could lump together the two alternatives (1.1) and (2.1) under the heading ' wholly *expressible* in non-ethical terms ' ; and the two alternatives (1.2) and (2.2) under the heading ' not wholly *expressible* in non-ethical terms '.

Now Frankena has argued that, if good were *simple*, it would be neither normative nor non-natural nor definable in terms of obligation. And we have interpreted (2c) to mean that, if good were neither normative nor non-natural nor definable in terms of obligation, there would be no reason to think that it is not definable in non-ethical terms. Putting the two together, we see that Frankena is committed to the proposition that, if good be *simple*, there is no reason to think that it is *not definable in non-ethical terms*. But, obviously, if it be simple, it cannot be *definable* in any terms whatever. So, in order to make sense of the above proposition, we must assume that Frankena is using the phrase ' not definable in non-ethical terms ' in a loose sense which is equivalent to my phrase ' not wholly expressible in non-ethical terms '. What he is asserting is, in fact, the following proposition. If good be neither normative nor non-natural nor definable in terms of obligation, there is no reason to think that it is not *either* (i) definable in wholly non-ethical terms, *or* (ii) identical with some simple admittedly non-ethical notion. If that is what Frankena means, it may be doubted whether (2c) is more than a tautology ; for the only attempt which he makes to define ' non-natural ' seems to identify it with ' not wholly expressible in certain enumerated non-ethical terms '.

Point 3. Frankena's third point may be put as follows. If good were either (*a*) *normative* or (*b*) *non-natural* or (*c*) *not wholly expressible in ethical terms*, then (α) it would be *definable in terms of obligation*, (β) it would *not be simple*, and (γ) it would *not be intrinsic*. The third point is therefore the conjunction of the nine hypothetical propositions which arise by uniting (*a*), (*b*), and (*c*) in turn as antecedents with (α), (β), and (γ) as consequents.

Now of these nine hypotheticals the following have already been dealt with. $(3a, \alpha)$ is the contrapositive of a step in the argument for point $(1, a)$, which Moore has discussed and dismissed. $(3a, \beta)$, $(3a, \gamma)$, $(3b, \beta)$, and $(3b, \gamma)$ are the contrapositives of $(2a, \alpha)$, $(2b, \alpha)$, $(2a, \beta)$, and $(2b, \beta)$ respectively. And these have been discussed and rejected by Moore. Again, if $(3b, \alpha)$ be granted, then $(3c, \alpha)$ becomes superfluous. For it is admitted that, if good be not wholly expressible in non-ethical terms, it is non-natural. And, if this be combined with $(3b, \alpha)$, we can infer $(3c, \alpha)$. $(3b, \alpha)$ embodies Frankena's conviction, already discussed, that the only fundamentally ethical notion is that of obligation. We are thus left with only $(3c, \beta)$ and $(3c, \gamma)$. These are more simply expressed in the equivalent form of their contrapositives. If we do this, and combine them, they amount to the proposition that, if good were either *simple* or *intrinsic*, it would be *wholly expressible in non-ethical terms, i.e.* it would be *natural*. This will best be treated incidentally in connection with the remaining points in Frankena's summary.

Point 4. Frankena's fourth point is that Moore has given no adequate reason for rejecting the view that ' good ' is *definable in terms of obligation*.

Moore begins by admitting that he has given no *conclusive* reason. But he thinks that he can give *good* reasons. The gist of his argument is as follows. He considers three alleged definitions of ' good ' in terms of obligation, which Frankena proposes. He rejects one of them on the ground that the two propositions suggested as definiens and definiendum do not even mutually entail each other. As regards the other two, he admits that there is mutual entailment between the definiendum and the suggested definiens. But he holds that that kind of logical relation can hold between two propositions without it being a case of two sentences with one and the same meaning. The test for the latter is to ask oneself the question ' Can I think of the one without *ipso facto* thinking of the other ? ' In each of these two cases he holds that that is possible, and therefore that there is not identity of meaning. Now, Moore says, he cannot think of any *other* plausible instances of mutual entailment between a value-proposition and an obligation-proposition. Therefore he holds that he has given sound, though not conclusive, reasons for thinking that goodness cannot be defined in terms of obligation.

I will now say something about the three proposed definitions. (1) ' X is intrinsically good ' means what is meant by ' If one is capable of producing X, one has a *prima facie* duty (in Ross's sense) to do so '. This is rejected by Moore (rightly, I think), on

the ground that the former proposition might easily be true when the latter was false.

(2) I am going to formulate the second in a slightly modified form of Moore's interpretation of Frankena's rather vague statement. It will run as follows. ' X is intrinsically good ' means what is meant by ' The mere fact (if it were a fact) that A could do Y and that Y would produce X would suffice to supply *some* reason for thinking that A *ought* to do Y '. Moore holds that these two propositions do entail each other. But he considers it obvious that a person could think of X as being intrinsically good, without *ipso facto* thinking of it as having this other complicated property, which is conveyed by and conveys its intrinsic goodness.

(3) Frankena quotes a certain alleged mutual entailment, which Moore gave in his *Ethics*, and asks why this should not be regarded as a definition. We need not trouble ourselves here about *this* particular alleged mutual entailment, because Moore says that he does not now think that it holds, or that the sentence quoted correctly expressed what he had in mind when he wrote his *Ethics*. Instead, we may confine our attention to the amended formula which he now proposes in its place. It runs as follows : ' X is intrinsically good ' entails and is entailed by ' If an agent were a Creator, before the existence of any world ; and if the only two alternatives open to him were (i) to create a world which consisted only of X, or (ii) to bring it about that there should never be a world at all ; then it would be his duty to choose alternative (i), provided (*a*) that he knew for certain that these were the only two alternatives open to him, and (*b*) that he did not think it wrong to choose alternative (i) '. Moore says that it seems obvious to him (and who shall deny it?) that a person could think of the former proposition without *ipso facto* thinking of the latter.

Point 5. The main assertion in Frankena's fifth point, and the only one which Moore discusses, is the following. Frankena alleges that, even though it be *intrinsic value* which makes a thing such that it *ought* to be pursued or brought into being by a competent agent, still Moore has given no good reason why ' intrinsic value ' might not be *definable in wholly non-ethical terms*.

What Frankena has in mind is no doubt this. He is alleging that there is no obvious reason why a *purely natural* characteristic, *e.g.* pleasantness, should not be such that the mere fact that a thing would have it would provide some ground for thinking that any agent, who could produce that thing, *ought* to do so.

29

In order to discuss this, let us begin by defining what Moore calls an 'ought-*implying* property', and what I prefer to call an 'ought-*inclining*' property. The sentence 'P is an ought-inclining property' is to mean what is meant by 'The mere fact that a thing would have P would suffice to provide *some* ground for thinking that any agent, who could bring such a thing into being, *ought* to do so'. Now Moore admits that *intrinsic goodness* is an ought-inclining property. He holds, moreover, that the intrinsic goodness of a thing always depends on the presence in it of some *natural* characteristic or other which is what I will call 'good-making'. Let Q be any *good-making* natural characteristic. Then anything that had Q would, of necessity, have intrinsic goodness. And the mere fact that anything had intrinsic goodness would suffice to provide some ground for thinking that any agent, who could bring such a thing into being, ought to do so. It follows at once that Q, though a *natural* characteristic, will also answer to the definition of an 'ought-inclining property'. And this can be generalized at once for every natural characteristic which is good-making.

There is no doubt, then, that Frankena is right in holding that there can be, and in fact are, *natural* characteristics which are ought-inclining. It is plain that he thinks that this fact entails that intrinsic goodness *either* (i) is *definable in terms of ought, or* (ii) is a *natural* characteristic. He thinks that, if the former alternative were fulfilled, there would be some reason to think that intrinsic goodness is *non-natural*. For, as we have seen, he regards 'ought' as *the* ethical notion *par excellence*, and as such the most plausible instance of a non-natural notion. On the other hand, he thinks that, if intrinsic goodness be *not* definable in terms of ought, then (in view of the fact that an ought-inclining property *can be natural*) there will be no valid reason for thinking that intrinsic goodness is non-natural.

Now Moore gives an argument which, he thinks, tends to show that intrinsic goodness cannot be identical with any natural property, even if it be not definable in terms of ought. The argument runs as follows.

Admittedly some *natural* characteristics are ought-inclining. But only *intrinsic* natural characteristics can be such. For a natural characteristic is ought-inclining only through being good-making. And only intrinsic natural characteristics convey intrinsic goodness. So the question reduces to whether intrinsic goodness could be identified with any *intrinsic* natural characteristic. After these preliminaries the argument continues as follows.

The number of ought-inclining intrinsic natural characteristics is, Moore asserts, certainly very great and possibly infinite. Plainly, we cannot identify intrinsic goodness with any particular one of them or with the aggregate of all of them.   Moore thinks it obvious, moreover, that intrinsic goodness could not be identified with the *disjunction* of all these natural characteristics. Suppose that there were *one single* non-disjunctive intrinsic natural characteristic, which was (*a*) ought-inclining, and (*b*) was conveyed by each of the other ought-inclining natural character-istics.   Then it might be plausible to identify intrinsic goodness with *it*.   But there seems to be no one natural characteristic answering to these conditions.   Therefore there does not appear to be any ought-inclining natural characteristic with which intrinsic goodness can be identified.   And it certainly cannot be identified with any intrinsic natural characteristic which is *not* ought-inclining.   Therefore it cannot be identified with any natural characteristic whatever.

I think that this argument is valid, so far as it goes.   But it would not satisfy a person who might suggest that 'X has intrinsic goodness' means what is meant by 'X has *some* intrinsic natural characteristic or other which is ought-inclining'.   I do not know whether Moore would count this as identifying intrinsic goodness with the *disjunction* of all ought-inclining intrinsic natural characteristics.   I do not think that he would.   But, if he did, I should be inclined to ask : What precisely is the objection to such an 'identification'?   The advantages of the suggestion are that it avoids postulating *two* indefinable non-natural charac-teristics, and defines the less specifically ethical one ('intrin-sically good') in terms of the more specifically ethical one ('ought').   The final objection would have to be that one can think of intrinsic goodness without *ipso facto* thinking of even so indeterminate a notion as that expressed by the phrase 'some intrinsic natural ought-inclining characteristic or other'.   But is that really at all certain?

Moore's latest published Views on the Connection of Good, Better, and Ought.   On pp. 606 to 611 of his terminal essay Moore formulates four very complicated pairs of mutually entailing propositions, which express the views, which he held at the date of writing, about the interconnections of 'good' and 'ought' and of 'better' and 'ought'.   I am going to state them in my own way ; but what I shall say is, I think, equivalent to what Moore had in mind and is perhaps somewhat easier to grasp.

I shall begin by defining certain statements.   (1.1) 'P is a *good-making* characteristic' means what is meant by 'If X did

have or now has or will have P, it follows that X then was or now is or will then be intrinsically good ; and if X should have had or should now have or should be going to have P, it would follow that X would have been or would now be or would be going to be intrinsically good '. (1.2) ' P is a *bad-making* characteristic ' is defined in a precisely similar way, with ' intrinsically *bad*' substituted throughout for ' intrinsically *good* '. (1.3) ' P is a *valifying* characteristic ' means what is meant by ' P is either a good-making or a bad-making characteristic '. (2) ' P is *more strongly good-making or less strongly bad-making* than Q ' means what is meant by ' If X did have or now has or will have P (and *no other* valifying characteristic), and if Y did have or now has or will have Q (and *no other* valifying characteristic), it follows that X then was or now is or will then be better than Y ; and if X should have had or should now have or should be going to have P (and *no other* valifying characteristic), and if Y should have had or should now have or should be going to have Q (and *no other* valifying characteristic), it would follow that X would have been or would now be or would be going to be intrinsically better than Y '.

We can now formulate the four pairs of mutually entailing propositions.

First Pair.   (i) P is a *good-making* characteristic.   (ii) If there had been, or in fact was, an agent who, before any world existed, (a) *knew* (α) that if he chose he could create a world characterized by P, (β) that he could so choose, and (γ) that if he did not so choose no world at all would ever exist ; and who (b) *did not believe* that this choice would be wrong ; then it would have been, or in fact was, the *duty* of that agent to make that choice.

Second Pair.   (i) P is a *good-making* characteristic.   (ii) P is an *ought-inclining* natural characteristic.

Third Pair.   (i) P is a *more strongly good-making or a less strongly bad-making* characteristic than Q.   (ii) If there had been or in fact was, an agent who, before any world existed, (a) *knew* (α) that if he chose he could create a world characterized by P, (β) that he could so choose, and (γ) that, unless he were so to choose, a world characterized by Q and not by P would inevitably come into existence ; and who (b) *did not believe* that that this choice would be wrong ; then it would have been, or in fact was, his *duty* to make that choice.

Fourth Pair.   (i) The world is *intrinsically better* because A chose to do Y, when he could have chosen to do something else instead, than it would have been if he had made any other choice open to him.   (ii) A *did his duty* in choosing Y.

Moore holds that in each of these four pairs the two members are interconnected by *synthetic* mutual entailment, but are not identical in meaning.   If either the first or the second or the third were *analytic*, it would provide a *definition* of intrinsic value in terms of obligation.   If the fourth were *analytic*, it would provide a definition of obligation in terms of intrinsic value.   For the reasons given, Moore does not regard any of them as analytic, and he therefore sees no reason to think that either notion can be defined in terms of the other.

(IV) *Ethical Egoism and Ethical Neutralism*

In my essay in the *G. E. Moore* volume I defined what I call 'Ethical Neutralism' and what I call 'Ethical Egoism'.   I pointed out that the latter might take milder or more extreme forms, but that it is in all its forms incompatible with Ethical Neutralism.   I thought that Moore had claimed in *Principia Ethica* (pp. 96 to 105) to show that Ethical Egoism (at any rate in its extreme form) is *self-contradictory*.   I argued that his attempt was a failure, and that all that could be proved was the tame proposition that Ethical Egoism is inconsistent with Ethical Neutralism.

Moore says, in his terminal essay, that what he was really trying to prove was not that Ethical Egoism *is self-contradictory*; but that Ethical Neutralism *would entail* that Ethical Egoism is self-contradictory, and not merely that it is false.   Now Ethical Neutralism is at any rate highly plausible, and to some eminent moralists it has seemed self-evident on inspection.   Therefore, if this argument of Moore's were acceptable, it would be at least highly plausible to hold that Ethical Egoism *is* self-contradictory, and not merely false.

Moore admits that his argument in *Principia Ethica* is extremely obscure and confused.   He now produces a new argument and it is with this that we shall be concerned.   It is extremely complex and hard to follow, and I am inclined to suspect that it contains a logical fallacy.   In order to try to show this as clearly as possible, I shall exhibit formally what I take to be Moore's new argument.

In what follows I shall write ' p *ent* q ' for ' p entails q ', and I shall understand by this a kind of logical relation which holds, *e.g.* between the conjunction of the premisses of a valid syllogism and its conclusion.   One way of describing it would be to say that the *conjunction* of p with not-q would be *impossible,* and that this impossibility does not depend on p being itself impossible

or on q being itself necessary. I shall write ' p *imp* q ' for
' p implies q ', and I shall understand by this that the conjunc-
tion of p with not-q is *in fact false*. With these notational
preliminaries, the argument may be stated as follows.

Let ' p ' stand for the sentence ' It would *not be wrong* for X to
choose Y '.

Let ' q ' stand for the sentence ' X does *not know* that the world
would be *intrinsically worse* if he were to choose Y than if he
were to choose some other alternative open to him at the time '.

Let ' r ' stand for the sentence ' X knows that the choice of Y
by him would procure *for himself* a more favourable balance of
intrinsically good over intrinsically bad experiences than any
other choice that he could make, and knows also that this choice
would be at least as favourable to the development of *his own*
nature and dispositions as any other that he could make '.

Then what Moore calls ' Proposition A ' is that p would follow
from r *alone*, even though q should be false. So we may write
' A ' for ' r *ent* p, even though not-q '. Moore asserts, and I
agree, that A is entailed by Ethical Egoism.

What Moore calls ' Proposition B ' is that the falsity of q
would entail the falsity of p. So we may write ' B ' for ' not-q
ent not-p '. Moore asserts, and I agree, that B is entailed by
Ethical Neutralism.

Now Moore asserts that A entails that r does *not* entail q. For,
he argues, to say that r would entail p, *even though q were false*,
entails that it is *logically possible* for q to be *false*, even though r
were *true*. This contention of Moore's may be written

$$A \text{ } ent \text{ not-}(r \text{ } ent \text{ } q)$$

Moore's argument may now be stated formally as follows :—

        B *ent* (not-q *ent* not-p)   (by definition)

Hence B *ent* (p *ent* q)   (by contraposition)   (I)

Again, A *ent* (r *ent* p)   (by definition)

Therefore (A & B) *ent* ((r *ent* p) & (p *ent* q))

Whence   (A & B) *ent* (r *ent* q)   (II)

But, as we have seen, according to Moore

$$A \text{ } ent \text{ not-}(r \text{ } ent \text{ } q)$$

Therefore (r *ent* q) *ent* not-A   (by contraposition)   (III)

Therefore ((II) & (III)) *ent* ((A & B) *ent* not-A)

Since (II) and (III) can be asserted, we can drop them and assert
what they together entail, *i.e.*

$$(A \text{ \& } B) \text{ } ent \text{ not-A} \qquad (IV)$$

Now up to this point the argument is valid, if we grant Moore's
contention (which I shall not here question) that A *ent* not-(r *ent* q).

But what he claims to have proved is that B entails that A is *self-contradictory*.   Now this must be the proposition

$$B \ ent \ (A \ ent \ not\text{-}A) \qquad\qquad (V)$$

(It must be clearly understood that it is *not* enough for Moore to show that

$$B \ ent \ (A \ imp \ not\text{-}A)$$

For A *imp* not-A is simply equivalent to (not-A or not-A), which is in turn simply equivalent to not-A.   So the latter proposition would merely amount to the tame conclusion that B *ent* not-A, *i.e.* in effect, that Ethical Neutralism is incompatible with Ethical Egoism.)

The question is, therefore, whether it is justifiable to infer from (IV), *i.e.* from (A & B) *ent* not-A, to (V), *i.e.* to B *ent* (A *ent* not-A). The answer is that this is *not* justifiable.   Consider, *e.g.* a valid syllogism (P & Q) *ent* R.   Suppose that you could legitimately derive from this the proposition P *ent* (Q *ent* R).   Then, if the premiss P were known to be true, you could drop it and assert the proposition Q *ent* R.   That this is not justifiable can easily be seen by taking a concrete example of a valid syllogism with a premiss known to be true.   Take, *e.g.* ' All men are mortal ' for ' P ', ' Socrates is a man ' for ' Q ', and ' Socrates is mortal ' for ' R '.   Then, if this kind of inference were valid, we could infer from the syllogism the proposition (All men are mortal) *ent* ((Socrates is a man) *ent* (Socrates is mortal)).   Then, dropping the true premiss that all men are mortal, we could assert that Socrates is a man *entails* that Socrates is mortal.   Now that conclusion is certainly false.   The mortality of Socrates is not a *necessary consequence* of his humanity *alone*.

So, unless I am much mistaken, Moore's new argument is fallacious, and he has failed to show that, if Ethical Neutralism were *true*, Ethical Egoism would be *self-contradictory*.   It is a rash undertaking to accuse Moore of a logical fallacy, and it may well be that I have misunderstood his argument.   On the other hand, it is very easy for the best of us to commit fallacies in *modal* logic, and so even Moore may have done so.   But that consideration cuts both ways, and I myself may have committed some fallacy in modal logic in my criticism of his argument.

University of Cambridge

II.—KNOWING AND BEING

By MICHAEL POLANYI

A FEW years ago a distinguished psychiatrist demonstrated to his students a patient who was having a mild fit of some kind. Later the class discussed the question whether this had been an epileptic or a hystero-epileptic seizure. The matter was finally decided by the psychiatrist: " Gentlemen ", he said, " you have seen a true epileptic seizure. I cannot tell you how to recognize it ; you will learn this by more extensive experience."

Clinical practitioners call the peculiar indescribable appearance of a pathological condition its *facies* ; I shall call it a " physiognomy ", so as to relate it to the delicately varied expressions of the human face which we can likewise identify without being able to tell quite how we recognize them. We may describe as a physiognomy also the peculiar appearance of a species which can be recognized only " aesthetically " [1] and further include among physiognomies the characteristics of wines and blends of tea which only experts can recognize.

Any effort to define a physiognomy more explicitly will aim: (1) at identifying its particulars and (2) at describing the relation between these particulars. I shall exemplify these two complementary efforts by reference to some comprehensive entities where the two can be observed separately.

(1) At the time when flying by aeroplane was first developed, around 1914-18, traces of prehistoric settlements were discovered from the air in fields over which many generations had walked without noticing them. Though the aerial photographs clearly revealed the outlines of the sites, the markings on the ground which constituted these outlines frequently remained unrecognizable. Such sites are comprehensive entities that are precisely traceable without mental effort from a distance, while the identification of their particulars at close quarters presents great difficulties. The unspecifiability of particulars is here as complete as it was for a hystero-epileptic fit, but with the difference that the relation of the unspecifiable particulars is clearly definable, which it is not in case of a physiognomy.

Usually it is not impossible to identify some particulars of a comprehensive entity, for example some symptoms of a clinically diagnosed disease. But in such cases another limitation of specifiability becomes apparent, as Gestalt-psychology has amply taught

[1] " Aesthetic recognition " is illustrated by C. F. A. Pantin, *Science Progress*, 42 (1954), 412.

458

us. Specifiability remains incomplete in two ways. First, there is always a residue of particulars left unspecified ; and second, even when particulars can be identified, isolation changes their appearance to some extent. Seen in isolation a jumping cat may appear to change its colour from white to dark grey, but when a window is included in the field of vision the cat is seen unchangingly the same white or grey. It has happened in Lancashire that a customer who ordered a shirting of green stripes on a white ground refused delivery, for he found the ground was purple. Since the colour of any patch of a surface varies with the context in which it is placed, coloured patterns are not specifiable in terms of their isolated particulars. The " red patch " generally used as a paradigm of a primary sense datum can never be unambiguously identified as such.

(2) Now for the reverse effort which aims at specifying the relations of particulars within a comprehensive entity. It is conspicuous in the case of topographic anatomy. We can easily identify the several organs, including blood vessels and nerve tracts in the human body (which is the task of systematic anatomy), but their mutual relation inside the body can be grasped only by a sustained effort of the imagination, based on the partial aspects revealed by successive stages of dissection. Similar difficulties have to be overcome in achieving an understanding of a geological topography and of the spatial relation of the component parts of a complex machine. Even the regular structure of crystals may present such problems. In 1923 H. Mark and I established the atomic structure of white tin. Shortly after that we had a visit by Professor van Arkel from Holland who claimed to have established an entirely different structure. Eventually, it transpired that this structure had the same arrangement of atoms as ours, but that he had described it along lines forming an angle of 45° to those along which we had seen it. This trivial difference in viewing the atomic arrangement had rendered it mutually unrecognizable to both parties, simply because we lacked a sufficient understanding of the relationships involved in the atomic arrangement. To this extent any complex spatial arrangement of opaque objects is unspecifiable.

We can see then two complementary efforts aiming at the elucidation of a comprehensive entity. One proceeds from a recognition of a whole towards an identification of its particulars, the other from the recognition of a group of presumed particulars towards the grasping of their relation in the whole.

I have called these two efforts complementary since they contribute jointly to the same final achievement, yet it is also true

that each counteracts the other to some extent at every consecutive step. Every time we concentrate our attention on the particulars of a comprehensive entity, our sense of its coherent existence is temporarily weakened ; and every time we move in the opposite direction towards a fuller awareness of the whole, the particulars tend to become submerged in the whole. The concerted advantage of the two processes arises from the fact that normally every dismemberment of a whole adds more to its understanding than is lost through the concurrent weakening of its comprehensive features, and again each new integration of the particulars adds more to our understanding of them than it damages our understanding by somewhat effacing their identity. Thus an alteration of analysis and integration leads progressively to an ever deeper understanding of a comprehensive entity.

Medicine offers readily an illustration of what I have in mind here. A medical student deepens his knowledge of a disease by learning a list of its symptoms with all their variations, but only clinical practice can teach him to integrate the clues observed on an individual patient to form a correct diagnosis of his illness, rather than an erroneous diagnosis which is often more plausible.

There is a close analogy between the *elucidation of a comprehensive object* and the mastering of a skill. A skill too is improved by alternate dismemberment and integration. Sportsmen like skaters, skiers, runners, swimmers ; artists like dancers, pianists, or painters ; skilled craftsmen and practitioners, all profit from motion-studies, followed by a skilful incorporation of the isolated motions into a complete performance.

Moreover, there are limits to the specifiability of skills and these are analogous to those which apply to the specification of physiognomies and other comprehensive entities. The analysis of a skilful feat in terms of its constituent motions remains always incomplete. There are notorious cases, like the distinctive " touch " of a pianist, in which the analysis of a skill has long been debated inconclusively ; and common experience shows that no skill can be acquired by learning its constituent motions separately. Moreover, here too isolation modifies the particulars : their dynamic quality is lost. Indeed, the identification of the constituent motions of a skill tends to paralyse its performance. Only by turning our attention away from the particulars and towards their joint purpose, can we restore in the isolated motions the qualities required for achieving their purpose. Again, as in the case of knowing a physiognomy, this act of integration is itself unspecifiable. Imitation offers guidance to it, but in the last resort we must rely on discovering for ourselves the right

feel of a skilful feat. We alone can catch the knack of it ; no teacher can do this for us.

The structural kinship of the arts of knowing and doing is indeed such that they are rarely exercised in isolation ; we usually meet a blend of the two. Medical diagnosis combines them about equally. To percuss a lung is as much a muscular feat as a delicate discrimination of the sounds thus elicited. The palpation of a spleen or a kidney combines a skilful kneading of the region with a trained sense for the peculiar feeling of the organ's resistance. It is apposite therefore to include skilful feats among comprehensive entities. Though we may prefer to speak of *understanding* a comprehensive object or situation and of *mastering* a skill, we do use the two words nearly as synonyms. Actually, we speak equally of *grasping* a subject or an art.

A peculiar combination of skilful doing and knowing is present in the working of our sense organs. I have mentioned before the interaction of all parts of the visual field in determining what we see. But it is our own active adjustment of our pupils and lenses and of the convergence of our eyes, that fashion the two retinal images on which this picture is based ; and the perceived picture depends on these actions once more as the messages received from the muscles adjusting the eyes are incorporated by us in the qualities of the perceived object. Nor is this all ; our perception is effectively co-determined also by messages from the internal ear, from the muscles which keep our body and head in its position, as well as by an ample range of memories. These internal data both guide the reflexes of our eye-muscles in shaping our retinal images, and control our evaluation in terms of perceptions of the sum total of relevant stimuli. There is also clear evidence to show that—as might be expected in the case of a skill—the capacity to see objects is acquired by training.

Visual perception does present, therefore, the main characteristics of a combined skilful doing and knowing. But there is also something new here, namely the fact that a major part of the particulars shaping the sight of an external object are *internal* actions and stimuli. Though our eye-muscles are subject to intentional control and we might even be dimly aware of a strain in them, such internal particulars are never clearly observable in themselves. And this unspecifiability is not of the same kind as found in the traces of a site which can only be seen jointly from the air ; for the particulars of perception (other than smell and taste) have the distinctive peculiarity of being projected from the interior of the body into the space outside it.

A similar projection takes place in the use of tools and probes, and the process can be studied here more easily, since the stimuli that are projected here can be fairly well observed in themselves. The relevant facts are well known. The rower pulling an oar feels the resistance of the water ; when using a paper-knife we feel the blade cutting the pages. The actual impact of the tool on our palm and fingers is unspecifiable in the same sense in which the muscular acts composing a skilful performance are unspecifiable ; we are aware of them in terms of the tool's action on its object, that is in the comprehensive entity into which we integrate them. But the impacts of a tool on our hands are integrated in a way similar to that by which internal stimuli are integrated to form our perceptions : the integrated stimuli are noticed at a distance removed outward from the point where they impinge on us. In this sense impacts of a tool on our hands function as internal stimuli, and a tool functions accordingly as an extension of our hands. The same is true of a probe used for exploring a cavity, or for a stick by which a blind man feels his way. The impact made by a probe or stick on our fingers is felt at the tip of the probe or stick, where it hits on objects outside, and in this sense the probe or stick is an extension of our fingers that grasp it.

The assimilation of a tool, a stick or a probe to our body is achieved gradually, as its proper use is being learned and perfected. The more fully we master the use of an instrument, the more precisely and discriminatingly will we localize at the farther end of it the stimuli impinging on our body while grasping and handling the instrument. This corresponds to the way we learn skilfully to use our eyes to see external objects.

We have now established analogous structures in several processes of knowing : namely, (1) the understanding of physiognomies, (2) the performance of skills, (3) the proper use of sensory organs, and (4) the mastery of tools and probes. In order to deal jointly with this whole family we need a general terminology for indicating the relation of a set of particulars to a comprehensive entity. The essential feature is throughout the fact that *particulars can be noticed in two different ways.* We can be aware of them uncomprehendingly, *i.e.* in themselves ; or understandingly, in their participation in a comprehensive entity. In the first case we focus our attention on the isolated particulars, in the second our attention is directed beyond them to the entity to which they contribute. In the first case therefore we may say that we are aware of the particulars *focally*, in the second, that we notice them *subsidiarily in terms of their participation in a whole.*

These alternatives do not necessarily involve any change in the degree of attention given to the particulars. When, after having first looked uncomprehendingly at the symptoms of a patient, we hit on the diagnosis of his illness, his symptoms become meaningful without becoming less noticeable. Focal and subsidiary awareness are definitely *not two degrees* of attention but *two kinds* of attention given to the *same* particulars. In the case of visual attention we may speak of *looking at* the particulars in themselves, as distinct from *seeing them* while looking at the context of which they form part. But " seeing " and " looking at " cannot be generally used instead of subsidiary and focal noticing.

We can formulate this difference in terms of *meaning*. When we focus on a set of particulars uncomprehendingly, they are relatively meaningless, compared with their significance when noticed subsidiarily within the comprehensive entity to which they contribute. As a result we have two kinds of meaning : *one* exemplified by the particulars of a physiognomy, where the two things, namely (*a*) the isolated particulars and (*b*) what they jointly mean, are not clearly separated in space ; and the *other*, exemplified by visual perception and the use of tools and probes, where the uncomprehended particulars are inside our body or at its surface, and what they mean extends into space outside. It may sometimes prove convenient to call the first kind of meaning *physiognostic* and the second *telegnostic*.

Our examples of unspecifiability can now be recapitulated in these terms. When a prehistoric site is discovered from the air we see the meaning of particulars which we cannot see in themselves. On the other hand, the study of topographic anatomy starts by seeing particulars in themselves uncomprehendingly, and then proceeds to realize their spatial relation, which represents their meaning, in the present sense of the term, *e.g.* to a surgeon. In the first case unspecifiability *impedes the analysis of a given meaning*, in the second case it *restricts the discovery of an unknown meaning*. In either case the alternation of analysis and integration progressively deepens both our insight into the meaning of a comprehensive entity in terms of its particulars and the meaning of these particulars in terms of their joint significance. When applied to the performance of a skill, this alternation renders our muscular actions ever better adjusted to their joint purpose.

We possess a supremely important system of telegnostic meaning in the denotative meaning of words. A man's name, by itself a meaningless sound, acquires a meaning by being consistently used as a pointer to the person whom it designates ; exactly

as the knocking of objects by a stick acquires a meaning as we learn to use the stick skilfully for feeling our way in the dark. The use of *general* denotative terms involves the establishment of a physiognostic meaning, namely the joint meaning of all the instances to which the general term is intended to refer. We meet meanings of analogous structure in the mastery of a skill, in the resourcefulness with which a master handles every new and unprecedented situation.

General conceptions claim a kind of foreknowledge which, though never quite absent from the act of knowing, is more marked in this case than in any other surveyed so far. This offers a convenient transition to the process of discovery in which foreknowledge will be seen manifested even more effectively. I shall lead on to this fact by showing first that all manner of discovery proceeds by a see-saw of analysis and integration similar to that by which our understanding of a comprehensive entity is progressively deepened. The two complementary movements are here a search for the joint meaning of a set of particulars, alternating with a search for the specification of their hitherto uncomprehended meaning in terms of yet unknown particulars.

The process of empirical induction may appear to present only the first of these two movements, since it proceeds from the observation of particulars to the discovery of their relations fixed by universal laws. But actually both movements are involved, for advances towards generalization do alternate with verification. The process of inductive discovery is in fact an oscillation between movements of analysis and integration in which, on balance, integration predominates.

Physics is largely based on such integrative discoveries, but in the progress of biology it is mostly analysis that leads the way. Discoveries aiming at a better understanding of living beings accept for their starting point the existence of plants and animals as complex entities, and usually proceed to analyse these in terms of their organs and the functions of their organs, which are analysed in their turn in terms of physics and chemistry. In biology these particulars are specified as means to an end and this applies also to the physical and chemical study of living beings. Physico-chemical processes may enable plants and animals to function well or else may disturb their functions, and they form the subject matter of biology only in so far as they have a bearing on life in one of these two ways.

But biological analysis alternates of course with integration and this may even predominate. New comprehensive entities

may be established in the manner in which Harvey discovered the circulation of the blood, partly by the observation of valves in veins, or when the process of evolution was established partly from palaeozoic data.

We must remember also that great advances have been achieved in physics by pure deduction whenever surprising implications of accepted theories were discovered. Indeed, to an important degree all discovery is deductive. For no enquiry can succeed unless it starts from a true, or at least partly true, conception of the nature of things. Such foreknowledge is indispensable and all discovery is but a step towards the verification of such fore-knowledge.

The structural analogy between knowledge and skill allows us to expand our perspective from discovery to invention. Fore-knowledge which guides discovery has its counterpart in the purpose which guides the invention of a skill, or of a new practical device. The most precious gift of the inventor is his capacity to conceive a practical purpose which might prove feasible. This is mostly the sole inventive function of a research director, who entrusts to his assistants the solution of the general problems set, or approved, by him. Agog with his problem, the inventor speculates on the possibilities offered by the field of experience, and by his sustained efforts to solve his problem brings about the emergence of its solution. Such is the heur-istic power of a practicable problem. It has been beautifully described on an elementary level by Koehler for chimpanzees. He shows how a desire pondered upon by the animal evokes a re-organization of its field of vision, revealing to it both the instru-ment and the procedure for satisfying its desire. This is also how skills are required and perfected ; this how we learn to use tools or probes ; and how we improve our powers of sensory per-ception. We may add that this is likewise the basic mechanism for solving a mathematical problem. The solution of a problem in mathematics is discovered by casting around with a view to filling the gaps left open in a situation which constitutes a solvable problem—exactly as a practical problem is solved by an invention.

We can lend greater precision now to what has been said earlier about the way induction is guided by our general concep-tion of the nature of things. Successful induction can be con-ducted only in the light of a genuine problem. An inductive problem is an intimation of coherence among hitherto uncompre-hended particulars and the problem is genuine to the extent to which this intimation is true. Such a surmise vaguely anticipates

the evidence which will support it and guides the mind engrossed by it to the discovery of this evidence. This usually proceeds stepwise, the original problem and surmise being modified and corrected by each new piece of evidence ; a process which is repeated until eventually some generalization is accepted as final.

To hit upon a problem is the first step to any discovery and indeed to any creative act. To see a problem is to see something hidden that may yet be accessible. The knowledge of a problem is, therefore, like the knowing of unspecifiables, a knowing of more than you can tell. But our awareness of unspecifiable things, whether of particulars or of the coherence of particulars, is intensified here to an exciting intimation of their hidden presence. It is an engrossing possession of incipient knowledge which passionately strives to validate itself. Such is the heuristic power of a problem.

But we may yet say that what is usually called knowledge is structurally similar to the knowledge of a problem. Knowledge is an activity which would be better described as a process of knowing. Indeed, as the scientist goes on enquiring into yet uncomprehended experiences, so do those who accept his discoveries as established knowledge keep applying this to ever changing situations, and to develop it each time a step further. Research is an intensely dynamic enquiring, while knowledge is a more quiet research. Both are ever on the move according to similar principles towards a deeper understanding of what is already known.

A theory of knowledge must be applicable to both kinds of knowing. Its function must be to justify our reliance on our knowing (whether dynamic or quiet) in spite of its unspecifiable contents. I shall outline this matter by using as my example the recognition of a disease by its physiognomy, which will have also to stand for all the other cases of knowing already surveyed.

The knowing of a disease is doubly unspecifiable. (1) We cannot identify, let alone describe, a great number of the particulars which we are in fact noticing when we diagnose a case of the disease. (2) Though we can identify a case of the disease by its typical appearance, we cannot describe it adequately, and there are four closely related reasons for this. (a) We are ignorant (according to 1) of the unspecifiable particulars which would enter into the description. (b) The relation between the particulars— even if they could all be identified—could be described only in vague terms which the expert alone would understand. (c) Our identification of a disease in any one instance comprehends

unspecifiably as its particulars the whole range of cases which, in spite of their individual differences, we have identified in the past, and (d) it relies on this comprehension for the future identification of an unlimited number of further cases which might differ from those known before in an infinite variety of unexpected ways. This is its heuristic function.

How can we justify such knowing ? Clearly not in terms of its unspecifiable contents ranging over (1) and (2), (a), (b), (c), (d). Any justification of it must credit ourselves with unformalizable powers of the kind that Kant acknowledged with reference to subsumptions of the type (2c) or (2d), by calling these powers " an art hidden in the depth of the human soul ". We must accredit in particular our competence for comprehending unspecifiable entities, which will yet reveal themselves in the future in an unlimited number of unexpected ways.

This may seem absurd. To claim that we can know the unexpected may appear self-contradictory. It would indeed be self-contradictory if knowing included a capacity to specify what we know. But if all knowledge is fundamentally tacit, as it is if it rests on our subsidiary awareness of particulars in terms of a comprehensive entity, then our knowledge may include far more than we can tell. The true meaning of the heliocentric system was discovered only by Newton, but it was anticipated 140 years earlier by Copernicus.

We can account for this capacity of ours to know more than we can tell if we believe in the presence of an external reality with which we can establish contact. This I do. I declare myself committed to the belief in an external reality gradually accessible to knowing, and I regard all true understanding as an intimation of such a reality which, being real, may yet reveal itself to our deepened understanding in an indefinite range of unexpected manifestations. I accept the obligation to search for the truth through my own intimations of reality, knowing that there is, and can be, no strict rule by which my conclusions can be justified. My reference to reality legitimates my acts of unspecifiable knowing, even while it duly keeps the exercise of such acts within the bounds of a rational objectivity. For a claim to have made contact with reality necessarily legislates both for myself and others with universal intent.

I must admit that I can fulfil my obligation to serve the truth only to the extent of my natural abilities as developed by my education. No one can transcend his formative milieu very far and beyond this area he must rely on it uncritically. I consider that this matrix of my thought determines my personal calling.

It both offers me my opportunity for seeking the truth, and limits my responsibility for arriving at my own conclusions.

This conception of knowledge as personal knowing departs in two closely related respects from the ideal of a strictly justifiable knowledge. It accredits man's capacity to aquire knowledge even though he cannot specify the grounds of his knowing, and it accepts the fact that his knowing is exercised within an accidentally given framework that is largely unspecifiable. These two acceptances are correlated within the effort of integration which achieves knowing. For this effort subsidiarily relies on the one hand on stimuli coming from outside, from all parts of our body and from tools or instruments assimilated to our body, and on the other hand on a wide range of linguistic pointers which bring to bear our pre-conceptions—based on past experiences—on the interpretation of our subject matter. The structure of knowing, revealed by the limits of specifiability, thus fuses our subsidiary awareness of the particulars belonging to our subject matter, with the cultural background of our knowing.

To this extent knowing is an indwelling : that is, a utilization of a framework for unfolding our understanding in accordance with the indications and standards imposed by the framework. But any particular indwelling is a particular form of mental existence. If an act of knowing affects our choice between alternative frameworks, or modifies the framework in which we dwell, it involves a change in our way of being. Since such existential choices would be included in an act of knowing, they could be exercised competently, with universal intent. Nor do similar existential changes, undergone passively, impair the rationality of our personal judgment. They merely affect our calling. For while they modify our opportunities for seeking the truth, they still leave us free to reach our own conclusions within the limits granted by these opportunities. All thought is incarnate ; it lives by the body and by the favour of society. But it is not thought unless it strives for truth, which leaves it free to act on its own responsibility, with universal intent.

We have found that our subsidiary awareness of the particulars of a comprehensive entity is fused, in our knowing of the entity, with our subsidiary awareness of our own bodily and cultural being. Let me now attend to the effect exercised by the subsidiary character of these particulars on the kind of being that we ascribe to the entity known through them.

When we watch a man's face and try to fathom his thoughts we do not examine his several features in isolation, but view them jointly as parts of his physiognomy. Thus we are aware

of far more particulars, and relations between particulars, than we could specify. Moreover, even the particulars which we could identify would lose their meaning when seen in isolation and would get lost among the irrelevant details of their surroundings. It is generally impossible, therefore, to keep track of a man's mental manifestations, except by watching them as pointers to the mind from which they originate. This is analogously true of any aggregate of unspecifiable particulars forming a comprehensive entity. But it is more pronounced in the traces of a human mind than it is, say, in the traces of a prehistoric site, because a human mind is actively at work in its own traces.

To show this I must go back to glance at the lower levels of life. The moment we identify a plant or animal we attribute an achievement to it. For we recognize it by its typical shape, which it has achieved by growing up healthily. At the same time we will notice any imperfections of its shape. Thus even when considering merely their shapes, living beings can be identified only in terms that attribute success or failure to them as individuals. On this morphological level the centre of individuality is comparatively weak. But the manifestations of this centre become steadily more accentuated as we successively ascend first to the vegetative level of physiological functions, then to active, sentient and appetitive behaviour, thence to the level of intelligence and inventiveness, finally reaching the level of the responsible human person. Each time we identify the existence of an individual at one of these levels we thereby attribute to it a measure of success or failure. The moment we recognize a human being we attribute to him a measure of active, responsible intelligence. We know a normal human being as a person, and the particulars of his physiognomy gain a vivid significance by being known in terms of this person.

At whatever level we consider a living being, the centre of its individuality is real. For it is always something we ascertain by comprehending the coherence of largely unspecifiable particulars, and which we yet expect to reveal itself further by an indeterminate range of future manifestations. Thus the criteria of reality are fulfilled. The human mind in particular is real in this sense, and indeed more real than the lower centres of individuality. For an intelligent person's mind is expected to reveal itself over a far wider range of indeterminate manifestations, than the lower centres which control his growth or appetitive action.

The progression from lower to higher levels of individuality involves a fundamental change in the relation between the

observed individual and ourselves as observers. When comprehending the lower levels of life in an individual, we ascribe success or failure to it by standards which we deem appropriate to its kind. Throughout the usual compass of biology, we find this relation between the observer and his subject : he is always critical of it. But as our study, ascending to higher levels, reaches that of man, criticism becomes mutual ; the subject of our study now criticizes us, just as we criticize it. Nor is this the end of the progression : our subject may ascend still further and become our master. We then become apprenticed to our subject and learn to accept its criticism of ourselves.

We arrive here at a process of education that, as such, is akin to the acceptance of our calling and which indeed, forms part of it. For our cultural background is determined to a considerable extent by the influence of a limited number of men. Primitive cultures are transmitted orally by a few authoritative persons in each succeeding generation. Our modern, highly articulate culture flows largely from a small set of men whose works and deeds are revered and consulted for guidance. The knowing of these great men is an indwelling in the sense already defined. Our awareness of their works and deeds serves us (to repeat my definition) as a framework for unfolding our understanding in accordance with the indications and standards imposed by the framework.

This progressive transition is no mere sleight of hand ; it is firmly grounded in the co-extension of knowing and being. We can see this by descending gradually from our educational dwelling in the works of great men to the biological study of plants and animals at lower levels. At no stage do we cease to participate in the life of the individual under observation. For we comprehend a living being at all levels by our subsidiary awareness of its particulars. These particulars are never observed in themselves ; we *read* them as manifestations of an individual. We rely on them as pointers, as we rely on a probe or a written text, by making them parts of ourselves for reaching beyond them. Thus our understanding of living beings involves at all levels a measure of indwelling ; our interest in life is always convivial. There is no break therefore in passing from biology to the acceptance of our cultural calling in which we share the life of a human society, including the life of its ancestors, the authors of our cultural heritage.

University of Oxford

III.—THE RETRIBUTIVIST HITS BACK

By K. G. Armstrong

When Kingsley Amis, in his book *Lucky Jim*, wanted to sum up the intellectual outlook of Professor Welch's wife, whose actions throughout show her to be excessively conservative, stodgy, snobbish, authoritarian, and generally disagreeable, he spoke of " Her opposition to social services because they made people lazy, her attitude to ' so-called freedom in education ', her advocacy of retributive punishment, her fondness for reading what Englishwomen wrote about how Parisians thought and felt ". This is interesting, for nowhere else in the book was punishment so much as mentioned. Amis uses advocacy of the Retributive theory purely as a symptom of a rather unpleasant, and certainly outdated, attitude, a symptom which his readers would surely recognize ; and I think Amis was right—most of the people who read his book would understand the reference just as he intended it, would regard the theory as something which, if it is not dead already, certainly should be.

C. S. Lewis wrote an article [1] defending retributivism. He is not a man whom one would normally expect to have any difficulty getting material published, yet the article appeared in an Australian periodical because, he said, he could get no hearing for his view in England. He claimed it had become clear during the controversy over the death penalty that his fellow-countrymen almost universally adhered to a combination of the Deterrent and Reformatory theories of punishment. In the British philosophical world the position is apparently similar. J. D. Mabbott says : " In the theory of punishment, retribution has been defended by no philosopher of note [for over fifty years] except Bradley. Reform and deterrence are the theories accepted in principle and increasingly influential in practice." [2]

" And ", it may be said, " This is just as it should be in an age that claims to be enlightened. Retributive punishment is only a polite name for revenge ; it is vindictive, inhumane, barbarous, and immoral. Such an infliction of pain-for-pain's-sake harms the person who suffers the pain, the person who inflicts it, and the society which permits it ; everybody loses, which brings out its essential pointlessness. The only humane motive, the only

[1] 'The Humanitarian Theory of Punishment ', *Twentieth Century* (Aust.), March 1949.

[2] J. D. Mabbott, ' *Contemporary British Philosophy* ' (ed. H. D. Lewis), p. 289.

possible moral justification for punishment is to reform the criminal and/or to deter others from committing similar crimes. By
making the punishment of wrongdoers a moral duty, the Retributive theory removes the possibility of mercy. The only
people who today defend the Retributive theory are those who,
whether they know it or not, get pleasure and a feeling of virtue
from seeing others suffer, or those who have a hidden theological
axe to grind. In any case, the theory is not only morally
indefensible but completely inadequate in practice to determine
what penalty the criminal should suffer in each case. Finally,
the theory can be shown to be wrong by such simple facts of
language usage as, for instance, that it makes sense to say ' He
was punished for something he did not do ', because, *inter alia,*
the theory demands that to say a man was punished for a crime
logically necessitates that he committed it. Historically,
morally, and logically, the theory is discredited."

These, then, are some of the objections that have been urged
against the Retributive theory of punishment. In my view they
are all mistaken—either because they are based on confusions
about what the theory is, or else because they spring from
erroneous moral judgments. One charge, that the proponents
of the theory are sadists or have a vested religious interest in it,
I shall not deal with. For all I know it could even be true, but
it seems irrelevant to the present discussion ; philosophers, of all
people, should surely be above using the ploy of analysing a man's
motives instead of meeting his arguments. As for the objection
that the theory is out of date, *i.e.* unfashionable, this seems so
ludicrous philosophically that I would not have mentioned it at
all if it were not for the unfortunate fact that it is the most
common objection of all. I shall certainly not bother with it any
further, giving the reader credit for being free from what Belloc
called " The degrading slavery of being a child of one's times ".

My aim is not to present, much less to establish as correct, a
full-scale Retributive theory, or set of theories, of punishment.
I want firstly to sort out the issues in this confusing and confused
area, and secondly to meet some of the attempts to discredit
retributivism which are made either by painting it, as I myself
did in summary fashion above, as a repulsive doctrine which
could only be held by the morally insensitive, or else by reducing
it to a harmless platitude which we all accept (so that the conflict
between it and other theories of punishment turns out to have been
no more than a foolish mistake). Let us start the process of
clarification by deciding just what is meant by a ' Theory of
Punishment ', and what it would be for such a theory to be wrong.

I

It soon becomes clear when one studies the contexts in which the phrase appears that there is not just one type of theory being invariably referred to. Usually a theory sets out to resolve one or more of three main problems, although all too often the reader is not told which of them is being tackled. Indeed, one is often forced to the conclusion that the writers are not clear on the point themselves, which has led to a great deal of confusion and many false oppositions between theories which are not simply different but which are attempting to solve different problems.

PROBLEM 1 : *Definition*

The first problem is over the meaning of the word ' punishment ', and is thus a definitional or logical issue. We examine the way the word is used in ordinary language, the things to which it is applied, and try to produce some rule which covers all these cases and only these cases. Applying the same technique to many other words we would get a single unequivocal answer on which all users of the language would agree—in short, the definition of the term—but in the case of ' punishment ' there is no such universally acceptable answer, and so we may speak of a ' theory of punishment ' in the sense of a claim that a certain definition exactly marks out the correct use of the term.[1] For such a theory to be wrong would be for it to mark out some range of things or activities not in fact ordinarily referred to by the word ' punishment ', or else to include only part of the proper range and/or more than the proper range. In this latter case we would probably say that the definition proposed was ' too narrow ' or ' too wide ', but the theory itself we would say was *wrong*, because it claims that the proposed definition exactly fits normal usage.

PROBLEM 2 : *Moral justification of the practice*

The second problem that a theory of punishment may be trying to solve is ' What, if anything, is the moral justification of punishment as such ? ' Why should it be felt that this particular practice requires moral justification, when in the case of so many other practices—from warning to washing-up—we do not feel that the question even arises? Clearly because punishment involves the deliberate infliction of pain, *i.e.* distress of some sort,

[1] The claimed definition that has attracted most attention in recent years is that given by A. G. N. Flew in his article ' The Justification of Punishment ', *Philosophy*, October 1954.

normally against the wishes of the recipient, and this is something to which there is a *prima facie* moral objection, the overriding of which requires justification.

It is important to notice that the moral justification of a practice is not the same thing as its general point or purpose, except in the eyes of those who have travelled so far down the Utilitarian road that they never question the means if the end is desirable. Every human practice that is not utterly random or unconscious has some point, but not all have, and many do not need, a moral justification. An act or practice may have a very sound point indeed and still lack moral justification, *e.g.* torturing prisoners to get information, so that to say that the general aim of the practice of punishing criminals is, say, the protection of society is not *eo ipso* to produce a moral justification of the practice, unless we assume that Bentham was right all the way. There is an ambiguity in a phrase like ' The general justifying aim of punishment ',[1] between *u hy* we do it, and why it is morally permissible—if it is—for us to do it.

A theory dealing with the moral justification of punishment as such could be wrong in two ways : firstly the general moral theory on which it is based could be incorrect, and secondly the theory of punishment could be a misapplication of the general moral theory.

PROBLEM 3 : *Penalty-fixing*

The third problem is this : which method or system of determining penalties for crimes is best? A theory of punishment dealing with this problem might better be called a theory of punishments, or a theory of penalties. The point of view from which the advocated method is said to be best varies ; sometimes it is in the interest of society as a whole, sometimes of the criminal, sometimes of both. One thing these theories have in common is that they are not so much concerned with what is as with what should be the case. To be wrong, such a theory would have to be advocating a method of determining penalties which was actually *not* best, either because it was not best from the point of view considered, or because some other factor wh'ch ought, morally, to be primary had been overruled.

Of these three problems, then, one is definitional, one is concerned with ethics, and one is largely practical, but with important moral overtones. The last two are very commonly dealt with as one, but the distinction is important for reasons which will

[1] This is a key phrase in H. L. A. Hart's article, referred to below.

become apparent later in this article.   Given this division of the problems of punishment, and the answering theories, into three categories, what of the attacks on retributivism?

II

It is generally considered that certain phrases which crop up fairly regularly in ordinary discourse, such as the troublesome ' He was punished for something he did not do ', create at least a *prima facie* difficulty for Retributive theories, and a number of different solutions have been offered in recent years.[1]   Because it seems to me that all these solutions are more or less mistaken, I want to look again at this issue in the light of the above analysis.

In the case of which of the three problems could a fact of language usage help to establish or prove wrong a theory of punishment?   For the definitional problem it would clearly be relevant, and I shall later examine the significance of a typical phrase. But can such a language fact prove anything about a theory which deals with the second or third problem, or, more specifically, can it prove such a theory wrong?   I suggest that it can do so only indirectly.

Any attempt to solve the second or third problem must assume that the logically prior problem of defining the word ' punishment ' has already been solved.   Unless we know what punishment *is*, what it means to impose a penalty for a crime, we cannot even start to talk about what system of fixing penalties is best, no matter from what point of view we may be considering it. Similarly, we cannot decide what it is that morally justifies punishment until we know what it is we are trying to justify. If the definition of ' punishment ' that these theories had worked with were shown not to be in line with usage then they would be wrong in the sense that they would have turned out not to be theories of punishment at all, but rather to be theories of something else.   Incidentally, at this point we can already see where A. C. Ewing goes astray in his book *The Morality of Punishment*, for he starts with the assumption that, whatever it is, punishment is morally justified, and then rejects some definitions because he cannot agree that what they produce is morally

---

[1] See especially A. M. Quinton, ' On Punishment ', *Analysis*, June 1954, reprinted in P. Laslett (Ed.), ' Philosophy, Politics, and Society ' ; A. G. N. Flew, ' The Justification of Punishment ', *Philosophy*, October 1954 ; K. Baier, ' Is Punishment Retributive ? ', *Analysis*, December 1955 ; and C. K. Benn, ' An Approach to the Problems of Punishment ', *Philosophy*, October 1958.

justifiable (*e.g.* on p. 34). Clearly the logical order is first to decide what punishment is, *then* to decide whether this thing is morally justifiable or not.

But if the term ' punishment ' *had* been correctly defined, is it still possible for a fact about usage to prove a theory dealing with the second or third problem wrong? I have already mentioned what it would be for such theories to be wrong in themselves; what, then, is the relevance of word usage?

Take the third problem. Settling what system of fixing penalties is best, no matter from whose point of view it may be considered, is essentially an exercise in practical reasoning. It is hard to see how the fact that the word ' punishment ' is sometimes used in certain phrases could ever show that, say, the criminal was not in fact better off when his sentence had been fixed on such-and-such principles. Yet this is exactly what is required to demonstrate that a theory of punishment of the third type is wrong.

The position with theories of the second type is somewhat similar. The way ethical terms are used certainly can show that a general moral theory is incorrect, but this is not true of the way that non-ethical terms are used. Now ' punishment ' is not in itself an ethical term : ' punishment ', like all activity words, can occur in ethical propositions, but such propositions are not made ethical by virtue of *its* presence. Nor, if the general moral theory was correct, and, by hypothesis, the term ' punishment ' had been correctly defined, could the theory of punishment be shown to be a misapplication of the general moral theory by some fact about word usage. But to establish the truth of this last assertion we will have to make a short excursion into the field of Ethics.

When it has been settled what it is for an activity to be moral or good (general moral theory), we still have to decide whether each particular activity, in this case the activity of punishing, is a case of a moral or good activity. The method employed to decide this varies with the general moral theory, but it will turn out to be one of the following kinds of procedure :

(i) An appeal to intuition in the broadest sense. To decide whether a particular type of activity is good, a duty, what one ought to do, etc., one has simply to reflect (not ratiocinate) on it and one can just ' see ' the answer. (Moore, Ross, in fact the majority of recent theories.)

(ii) A factual calculation of the total amounts of pleasure and pain that the action causes. (Hedonistic Utilitarianism.)

(iii) (*a*) A check on whether God has told us, by Revelation, to do it. (The theory that Good is that which God enjoins.)

    (*b*) A check on whether the majority of the community approves of it. (The theory that Morality is social convention, *i.e.* Social Externalism.)

    (*c*) A simple statement of whether the speaker himself likes or approves of it (Subjectivism) and wants others to do so too (Stevenson, Emotivists generally).

(iv) Settling whether it is in accordance with Human Nature and Man's Final End, both by examining our internal, intuitive attitude to it and by reasoning from what we already know of Man's nature and destiny. (Thomist theory.)

(v) Checking it against a set of specific criteria of various sorts provided by the general theory for determining what is in accordance with the Moral Law (Kant) or what are genuine moral rules (Baier).

Now if we consider all these methods it can be seen that in no case could the theory of punishment (moral justification) produced by their use be upset by facts about the use of the word 'punishment'. Remember that the original data about what it is for an activity to be good, moral, etc., and what punishment *is* are, in each case, correct by hypothesis. In method (i) no further data at all are introduced, so there is no possibility of error through false information. In method (ii) the additional information is scientific, mainly psychological; it is about how men feel, not about how they use words. In method (iii) there is room for error over (*a*) what God *has* commanded, or (*b*) what the majority of the community *does* approve of—it is very doubtful whether one could make an error over (*c*) what oneself approves of or likes—but neither of these could be *shown* to be erroneous by the way the word 'punishment' is used in sentences not about Revelation or approval. In method (iv) an error could only come in through a false notion of Human Nature or mistakes about Man's Final End; but our ideas about Man's Nature and End are in no way dependent on the question of which sentences using the word 'punishment' make sense and which do not. In method (v) the possible sources of error will vary with the criteria put up by the general moral theory for determining whether an activity constitutes a breach of a genuine moral rule. However, the only criterion which could be shown to have been misapplied by our noting that it made sense to use the word 'punishment' in some given non-ethical sentence would be one which specified that this must not be the case, *e.g.* 'An activity, to be moral,

must be such that the word signifying it cannot sensibly be used in such-and-such sort of non-ethical sentences '. Now, of course, no general theory of morals which would lead to the use of method (v) has such a criterion, and it is hard to see what reason there could ever be for introducing such a one.

We can see, then, that even if a theory of the moral justification of punishment can be wrong in the sense of being a misapplication of the correct general moral theory, whichever that may be, its wrongness can never be proved by the fact that it makes sense to use the word ' punishment ' in some given nonethical sentence. The stage of appeal to language habits has already been passed.

So far, we have established that while a fact about how the word ' punishment ' is used might well show that a theory of punishment in the definitional sense was wrong, such a fact could not show that a theory dealing with the moral-justification or penalty-fixing problems had done so incorrectly, except in the sense that they had dealt with something other than punishment. We must now turn back to definitional theories of punishment, to see what precisely is the effect on them of the fact, if it is a fact, that it makes sense to say, for example, " He was punished for something he did not do ".

Irrespective of which problem or problems it sets out to solve, a theory of punishment can usually be put under one of three headings : Retributive, Deterrent, or Reformatory. When the problem is to define punishment these theories provide roughly the following answers :

1. *Retributive* : Punishment is the infliction of pain, by an appropriate authority, on a person because he is guilty of a crime, *i.e. for* a crime that he committed. I do not intend to go into the question of just what constitutes an appropriate authority, because the answer would appear in all three definitions, and it is the *differences* between them, concerning who suffers pain and why, that I wish to stress here. Also, I use the word ' crime ' deliberately, as it is ambiguous between an offence against a rule, the law, morality, or someone else's rights.[1]

2. *Deterrent* : Punishment is the infliction of pain on a person in order to deter him from repeating a crime or to deter others from imitating a crime which they believe him to have

[1] For discussions of what the offence is against, see J. D. Mabbott, MIND, April 1939 and *Philosophy*, July 1955, and C. H. Whiteley, *Philosophy*, April 1956.

committed.   I am here subscribing to Benn's view that deter-
rence of the person punished is not reform.   Reform means
that the man intends to avoid repeating the crime, not from
fear of punishment but because he sees that it was wrong.
(C. K. Benn, *op. cit.*)

3. *Reformatory* :   Punishment is the infliction of pain on a person
   in order to reduce his tendency to want to commit crimes or to
   commit crimes of a particular sort.

I am not urging the acceptance of any one of these three definitions
in preference to the others.   I have set them out so that we may
see what difficulties arise for retributivism in this area from the
alleged fact that it makes sense to say " He was punished for
something he did not do ".

What does it come to to say that X makes sense, where X is
some sequence of words?   If we take it loosely, in what I shall
henceforth call the weak way, it comes to saying that when we
hear X we can understand what is being asserted by the speaker,
and this does not necessarily imply that he is using all the words
contained in X in their exactly proper way, but only that what is
said is, so to speak, ' near enough '.   Thus, for instance, " They
half killed him " makes sense in the weak way, even to someone
not familiar with the phrase as an idiom, although killing,
strictly speaking, is a deed that allows of no degrees—a man is
either killed or he is not.   But we can also take ' X makes sense '
in a tighter way, which I shall henceforth call the strong way, as
indicating that each word has been used quite correctly so that an
analytical substitution can be made for any term involved
without revealing any contradiction, inconsistency, or other
logical impropriety.

If ' He was punished for something he did not do ' makes sense
in the weak way, how does this affect the Retributive definition?
It is fairly easy to see that it is not incompatible with the truth
of the theory.   The person who used the sentence could simply
be asserting that someone who was not in fact the perpetrator
of a particular crime had been treated as though he were, either
because those in authority held sincerely though mistakenly that
he *was* guilty of it or because they had deliberately tried to mis-
lead the public.

If, however, the given sentence made sense in the strong way,
it is equally obvious that the Retributive Theory (definition)
would be proved wrong.   Since, on the definition it proposes, to
say that someone was punished for a crime involves saying that
he committed it, the Retributive definition would make " He

was punished for something he did not do " a self-contradictory proposition. Thus *either* the sentence does not make sense in this strong way *or* the Retributive Theory (definition) must be abandoned.

But any satisfaction which advocates of the Deterrent or Reformatory Theories might feel over this incompatibility must be short lived, for if the sentence did make sense in the strong way this would be equally fatal to their own proposed definitions, although for a different reason. Here the difficulty lies not in " . . . that he did not do ", but in " . . . *for* something . . . ". It is clear that the ' something ' referred to is an act, a crime that somebody committed ; but on neither the Deterrent nor the Reformatory Theory is a man subjected to pain *for* a crime, but *to* deter him or others from committing crimes in the future or *to* rid him of the tendency to commit crime. The only theory with which the given sentence's making sense in the strong way would be compatible would be one which defined punishment as ' The infliction of pain on a person because a crime has been committed, whether by that person or not '. As far as I know, such a theory is not held by any philosopher in the Western World. That this should be so is, I suggest, strong *prima facie* evidence that the sentence does not in fact make sense in the strong way.

To recapitulate : Only theories of punishment of the definitional type could be affected by a fact of language not involving ethical terms. If it makes sense in the weak way to say " He was punished for something he did not do " then the Retributive theory (definition) is still tenable ; if it makes sense in the strong way then not only the Retributive, but also the Deterrent and Reformatory definitions are shown to be wrong. Thus the alleged difficulty is either no difficulty at all, or else it is an insuperable difficulty for all three definitional theories. In neither event are there grounds in word usage for discriminating against retributivism *vis-a-vis* its rivals.

In passing we may note an interesting asymmetry between the sentence we have been discussing and the sentence " He was punished although he was innocent ", where this is interpreted to mean not just that the man was innocent of the particular crime for which he was ' punished ' but that he was ' punished ' although he had not committed any crime at all. If " He was punished although he was innocent ", interpreted thus, makes sense in the strong way, then whilst the Retributive theory (definition) is shown to be wrong the Deterrent and Reformatory theories (definition) are still viable.

The treatments of this question given by Flew and Quinton have already been effectively criticized by Baier (*op. cit.*) ; but Baier's own analysis differs from the one I have just set out. In his view, the fact that it makes sense to say " He was punished for something he did not do " is *not* incompatible with the Retributive theory, because the theory necessitates as a logical precondition of punishment not that the person who suffers the pain *in fact* committed the crime, but only that he was ' *found guilty* ' of committing it. This ' finding guilty ' can either be informal and implicit (as in the case of a parent punishing a child) or formal and explicit (as in the case of a jury announcing its verdict). But the fault in this solution is that whilst the amended Retributive definition it is based on includes everything that we would call punishment, it also includes things we would *not* call punishment. Merely to go through the moves of the ' game ', as he puts it, is not enough to constitute a case of punishment. Take the case of a man who clearly did not commit a crime, and who is obviously not even believed to have committed a crime ; if, despite his known innocence, a court went through the formal motions of a trial, including the moves of the jury uttering the word " Guilty " at the appropriate time and the Judge sentencing him, and if he was duly executed—would it be proper to call this a case of punishment? Or consider a schoolmaster regularly beating one of his pupils after saying the words " Bloggs ! You were laughing ! " (an informal declaration of guilt) although Bloggs never so much as cracks a smile and the teacher's eyesight is good—would we say Bloggs was being *punished* regularly? Would it not be more natural to say that these were cases of something other than punishment—victimization perhaps?

III

Earlier, I claimed that a lot of contemporary writing on the subject of punishment is confused, much of the trouble springing from a failure to distinguish between the three separate problems involved. I can best support this claim by taking my examples from the two most quoted post-war articles on punishment.[1] Consider first this paragraph written by Flew [2] in criticism of J. D. Mabbott's article " Punishment " : " The objection to saying that the *sole* justification for punishing someone is that he

[1] A. G. N. Flew, "The Justification of Punishment", *Philosophy*, October 1954, and A. M. Quinton, "On Punishment", *Analysis*, June 1954.

[2] *Op. cit.*, p. 301. Mabbott's article appeared in MIND, April 1939.

has committed an offence is that Mabbott and almost everyone else would allow that a punishment in certain circumstances was overdetermined in its justification—was justified twice over. Certainly : because, though Mabbott claims to ' Reject absolutely all utilitarian considerations from its justification ', he is prepared to appeal to these to justify *systems* of punishment.   But if a *system* is to be justified even partly on such grounds, some cases within the system must be partly justifiable on the same grounds : the system surely could not have effects which no case within it contributed."

Here Flew is confusing Problems 2 and 3.   The point is that justifying *systems* of punishment, *i.e.* which method or system of determining penalties for crimes is best (Problem 3), is quite distinct from justifying, morally, the practice of punishment as such (Problem 2).   There is nothing inconsistent in Mabbott's view that, whilst the moral justification of inflicting pain on people who commit crimes lies solely in retributive considerations, one system of fixing *what* penalties are to be inflicted for *what* crimes may be better than another from the point of view of its consequences on a particular society, *i.e.* on ' utilitarian ' grounds. Flew has failed to realize that the word ' justification ' is used in two quite different ways in the context of discussions of punishment, depending on which problem a theory is trying to solve. When dealing with Problem 3 a *system* is ' justified ' precisely by establishing that it *is* most in the interests of some stated person or group that penalties should be fixed on the lines it prescribes. Of course it is always possible that some system of determining penalties which has been ' justified ' in this way may be objected to on *moral* grounds if it goes against a principle arising from the solution to problem 2, *e.g.* if a system of fixing penalties, ' justified ' by its deterrent effect alone, resulted in overruling the moral principle (derived from a Retributive solution to Problem 2) that a very minor offence ought not to be punished more severely than a very serious one.   But this is not surprising ; if I say that punishment has a moral justification, I do not thereby resign my right to apply moral criticism to any system of fixing penalties.   I might, for instance, hold that the practice of punishing is morally justifiable, yet at the same time say that the Nazi system of partly determining penalties according to the race of the criminal was immoral.

In his article " On Punishment " Quinton recognizes that theories of punishment may deal either with what punishment *is* or with the problem of morally justifying it as a practice (Problems 1 and 2).   However, when he proceeds further he confuses

the issue by misunderstanding retributivism in two ways. Firstly, he says that it is a logical and not a moral doctrine : " It does not provide a moral justification of the infliction of punishment but an elucidation of the use of the word " (p. 134). We have seen that this is true of the Retributive theory, as it is also true of the Deterrent and Reformatory theories, *but only when the theory deals with Problem* 1. If a particular Retributive theory deals with Problem 2, as it well may, then it *is* a moral doctrine, albeit one that Quinton does not agree with, *e.g.* " The moral justification of punishment is simply that the infliction of pain on those who have inflicted pain on others is a Good-in-itself, since it is a species of justice " would be a possible, though perhaps poor, Retributive theory dealing with Problem 2. Of course it is true that Retributive theories are very often concerned only with definition, but they can be, and sometimes are, concerned with moral justification or systems of penalty-fixing, *e.g.* C. S. Lewis's article ' The Humanitarian Theory of Punishment ' (*op. cit.*) argues for a Retributive theory of penalty-fixing, and A. C. Ewing (*op. cit.* Chap. 2.) argues against such a theory, thus recognizing its existence.[1] Secondly, Quinton misunderstands retributivism when he says that it regards punishment as trying to bring about " A state of affairs in which it is as if the wrongful act had never happened " (p. 135). He criticizes this doctrine as only applicable to a restricted class of cases : " Theft and fraud can be compensated, but not murder " (*loc. cit.*). Here he is confusing *retribution* with *restitution*. If we recover stolen property, or if a confidence man repays the money he got by fraud, then although restitution has been made the retributivist would say that punishment was still due, *i.e.* the *loss* has been annulled but the *crime* has not. Only physically are things as they were before the crime. In the case of murder, restitution is clearly impossible—we cannot get back the life that was taken—but Retributive punishment is still possible. Further, the *lex talionis* is not an *extension* of retributivism, as Quinton claims (*loc. cit*), but a particular Retributive theory dealing with Problem 3 (penalty-fixing), and in my view a poor Retributive theory, as I shall explain later.

[1] In a note published after this article was written, A. S. Kaufman points out that F. H. Bradley, at least, was a retributivist who was not only concerned with the definition of ' punishment ', but had a Retributive view on the moral issue. (' Anthony Quinton on Punishment ', *Analysis*, October 1959.)

31

IV

What of the other charges against retributivism?   Is it, as is so often said, inhumane?   This charge, if correct, would count as a moral objection against a Retributive theory of penalty-fixing (Problem 3).   In the area of this problem it seems to me that Retributive theories stand up very well to comparison with purely Deterrent or Reformatory theories.   If we penalize the criminal according to what he has done, we at least treat him like a man, like a responsible moral agent.   If we fix the penalty on a Deterrent principle (*i.e.* What penalty given to this criminal, or class of criminal, will effectively deter others from imitating his crime?) we are using him as a mere means to somebody else's end, and surely Kant was right when he objected to that !   And why stop at the minimum, why not be on the safe side and penalize him in some pretty spectacular way—wouldn't that be more likely to deter others?   Let him be whipped to death, publicly of course, for a parking offence ;   that would certainly deter *me* from parking on the spot reserved for the Vice-Chancellor !   And of course a deterrent will deter as long as the person on whom the pain is inflicted is *believed* to be guilty by those we wish to deter.   It really wouldn't matter, if deterrence is our aim in fixing penalties, whether he was in fact guilty or not ;   as long as we kept his innocence a secret we could make a very effective example of him.   This conclusion has been acted on by more than one government in our own times.

If, on the other hand, our aim in fixing penalties is the reform of the criminal—his *cure*, some might say—then the logical pattern of penalties will be for each criminal to be given reformatory treatment until he is sufficiently changed for the experts to certify him as reformed.   On this theory, every sentence ought to be indeterminate—' To be detained at the Psychologist's pleasure ', perhaps—for there is no longer any basis for the principle of a definite limit to punishment.   " You stole a loaf of bread?   Well, we'll have to reform you, even if it takes the rest of your life."   From the moment he is found guilty the criminal loses his rights as a human being quite as definitely as if he had been declared insane.   This is not a form of humanitarianism I care for.   Nor does it become any more humane if we drop the word ' punishment '—it is still just as compulsory.   C. S. Lewis wrote a sentence on this point that is worth quoting, even if only as a masterly piece of propaganda :   " To be taken without consent from my home and friends, to lose my liberty, to undergo all those assaults on my personality which modern psychotherapy

knows how to deliver, to be remade after some pattern of 'normality' hatched in a Viennese laboratory to which I never professed allegiance, to know that this process will never end until either my captors have succeeded or I have grown wise enough to cheat them with apparent success—who cares whether this is called punishment or not." (C. S. Lewis, *op. cit.*) And, since prevention is better than cure, why wait until he commits a crime? On the Reformatory theory of penalty-fixing, it is the *tendency* to commit crimes that we want to eliminate, so if a man has the tendency let him be penalized before the damage is done. Let him be penalized for what he is, not for what he does, and let him be made over into what the authorities (or their experts) want him to be.

The usual riposte to the sort of charges I have been making against Deterrent and Reformatory theories of punishment (penalty-fixing) is to refer back to a Retributive *definition* of punishment and rule out the charges as logically inadmissable.[1] " The short answer to the critics of Utilitarian theories of punishment ", writes C. K. Benn (*op. cit.* p. 332), " is that they are theories of *punishment*, not of any sort of technique involving suffering ". But to say, to those who ask why we shouldn't punish the innocent when it would be socially useful, that " The infliction of pain on a person is only properly described as punishment if that person is guilty " [2] is to give an answer which is technically correct (for those who subscribe to a Retributive definition of punishment) but which misses the point behind the question.

Suppose the questioner comes back as follows : " All right then, if you want to quibble about terminological niceties when I'm trying to make a serious moral and practical enquiry, I'll rephrase my question. Why shouldn't we do to the innocent that which, when it's done to the guilty, is known as punishment ? " At this point those theorists who offer a Utilitarian moral justification for the practice of punishing (*i.e.* a Deterrent and/or Reformatory theory on Problem 2) are in a difficult position ; for on their view what morally licenses us to inflict pain on a man is not that he is guilty—that is merely what gives us a *logical* licence to use the word ' punishment ' to refer to the infliction of pain—

---

[1] In an important and constructive paper, which did not appear until after this article was written, H. L. A. Hart has coined the name 'Definitional stop' for this sort of riposte. (' Prolegomenon to the Principles of Punishment ', The Presidential Address to the Aristotelian Society, 1959-60, p. 5.)

[2] A. M. Quinton, *op. cit.* p. 137.

but that there will be a socially useful result in terms of his reform and/or the deterrence of others, or, to put it more generally and in Mr. Benn's terminology, that the decrease in mischief to the public will be greater than the increase in mischief to those who are subjected to the pain.   The only objection these theorists could raise would be that inflicting pain on the innocent is not in fact an effective deterrent.   This empirical hypothesis is of very doubtful validity—it is not hard to think of cases where a very great mischief to the public might be avoided by condemning and executing an innocent man under guise of punishing him, *e.g.* who knows but that Klaus Fuchs might have been deterred from passing information on the A-bomb to the Russians if the Government had previously ' framed ' some innocent scientist on an espionage charge and executed him in a blaze of publicity? In any case, most people feel that there is more against ' punishing ' the innocent than that it wouldn't effectively reduce crime. Nor does our sense of outrage arise solely from the lying imputation of guilt, as Quinton claims (*loc. cit.*), although this is undoubtedly a partial explanation.   Surely our principal objection is to the deliberate infliction of *undeserved* pain, to the *injustice* of it, and this moral objection to taking Deterrent and Reformatory theories of penalty-fixing (Problem 3) to their logical conclusion can only be accounted for on a *Retributive* theory of the moral justification of punishment as such (Problem 2), as I shall show in a moment.

But before we leave the question of penalty-fixing it is worth asking why it should be so often thought that Retributive theories in this area are necessarily barbarous.   The charge springs from the misconception, which I mentioned before, that there is only one such theory—the *lex talionis*.   In fact, all that a Retributive theory of penalty-fixing needs to say to deserve the name is that there should be a proportion between the severity of the crime and the severity of the punishment.   It sets an upper limit to the punishment, suggests what is *due*.   But the ' repayment ' (so to speak) need not be in kind ; indeed in some cases it *could not* be.   What would the *lex talionis* prescribe for a blind man who blinded someone else?   Even in those cases where repayment in kind of violent crime is possible there is no reason why we should not substitute a more civilized equivalent punishment ;   the scale of equivalent punishments will, of course, vary from society to society.   There is also no reason, having got some idea of the permissible limits of a man's punishment from Retributive considerations, why we should not be guided in our choice of the form of the penalty by Deterrent and Reformatory considerations.

In the area of the moral justification of the practice (Problem 2) a Retributive theory is essential, because it is the only theory which connects punishment with desert, and so with justice, for only as a punishment is deserved or undeserved can it be just or unjust. What would a just *deterrent* be? The only sense we could give to it would be a punishment which was just from the Retributive point of view and which also, as a matter of fact, deterred other people. " But ", it may be objected, " You are only talking about *retributive* justice." To this I can only reply : What other sort of justice is there?

A vital point here is that justice gives the appropriate authority the *right* to punish offenders up to some limit, but one is not necessarily and invariably *obliged* to punish to the limit of justice. Similarly, if I lend a man money I have a right, in justice, to have it returned ; but if I choose not to take it back I have not done anything unjust. I cannot claim more than is owed to me but I am free to claim less, or even to claim nothing. For a variety of reasons (amongst them the hope of reforming the criminal) the appropriate authority may choose to punish a man less than it is entitled to, but it is never just to punish a man more than he deserves. It is a mistake to argue—as Ewing, for example, does in Chapter II of *The Morality of Punishment*—that, on the Retributive theory, to punish a man less than the exact amount due is an injustice similar to punishing an innocent man. The Retributive theory is not, therefore, incompatible with mercy. Quite the reverse is the case—it is only the Retributive idea that makes mercy *possible*, because to be merciful is to let someone off all or part of a penalty which he is recognized as having deserved.

Retributive punishment is not revenge, although both are species of justice. Revenge is private and personal, it requires no authority of one person or institution over another ; punishment requires a whole system of authorities given a right to secure justice. As members of the State, we surrender the right to secure justice ourselves to the authorities that the State appoints (though retaining, for example, our right to punish our own children). It is these State-appointed authorities, not ourselves, who must both punish malefactors and recover for us, by force where necessary, what a reluctant debtor owes us.

Finally, it is a distortion of the Retributive theory to say that it involves the infliction of pain-for-pain's-sake. On my understanding of the theory, pain (which the appropriate authority is morally licenced to inflict because it is deserved) is inflicted for the sake of all or any of a number of different ends. Amongst

these are the protection of society, the reform of the criminal—which punishment *may* achieve simply by making the criminal realize the full gravity of what he has done, as a child realizes how serious his offence has been when he sees how angry it makes his father—and the deterrence of others. All these ends are in themselves both morally and socially desirable; where the infliction of pain is justified by desert, pain may be a morally permissible *means* to achieving them.

V

I do not claim to have demonstrated in this article that Retributive theories are the correct solutions to each of the problems of punishment. My aim has been to make clear the distinction and interconnection between those problems and to show that, if we are to reject Retributive theories, objections more powerful than those currently advanced and accepted will have to be found, since the current objections rest on confusions, or mis-statements of the problem, or mere prejudice.

Where the problem is to define punishment (Problem 1), some sort of Retributive theory now seems to be fairly generally accepted, although whether the offence for which punishment is inflicted is against the Law, some other set of explicit rules, or just the accepted moral standards of a community is still debated. Most of those who have examined the difficulty raised by the currency of the phrase ' He was punished for something he did not do ' rightly conclude that the Retributive definition is more or less immune, even if their reasons for so concluding are faulty. That able philosophers should accept such unsatisfactory solutions to the apparent problem bears witness to the strength of their conviction that a Retributive definition is the right one.

The moral justification for the practice of punishment as such (Problem 2) is today sought almost invariably in Reformatory or Deterrent terms. For those who subscribe to simple Utilitarian theories of morals of the total-pain-and-pleasure-to-society type, this is of course automatic, and short of refuting their moral theory as a whole one cannot hope to shift their position on punishment. But most of those philosophers for whom retributivism is not ruled out *a priori* by their general moral theory— *i.e.* most of that great majority who do not subscribe to Hedonistic Utilitarianism—also reject or, more commonly, ignore the Retributive theory of the moral justification of punishment; and this despite the unique ability of that theory to connect punishment with the notions of desert and justice and, indeed, with the

deep-seated general conviction we all have that to strike back and to strike first are two very different things, morally speaking, irrespective of the results they may produce or be intended to produce.   This is surprising ; for though I have not demonstrated that these considerations in favour of a Retributive theory could never be outweighed by any conceivable arguments for Deterrent and/or Reformatory moral justification, I hope I *have* shown that they are too important to be altogether overlooked, or even to be summarily dismissed.   It may be possible not to be moved by them, but they must at least be faced.   I have referred to three factors which tend to account for the cavalier treatment of retributivism in the area of Problem 2.   First, there is the mistaken belief that a Retributive moral justification of punishment would make the infliction of pain on the guilty a positive, inescapable obligation, instead of merely creating a right to inflict pain which, like other rights, it may in some circumstances be foolish or mean to exercise.   Second, there is the failure to distinguish between the moral justification of a practice on the one hand, and its ' general justifying aim ' on the other.   And third, there is the notion that if one concedes the definitional field to retributivism, there is no further area in which Deterrent and Reformatory theories can have a Retributive rival, that one can ' dissolve ' the traditional conflict between the theories by declaring that ' Retributivism is not a moral but a logical doctrine '.

When the problem is to find the best system of penalty-fixing there is no doubt that a purely Retributive theory would have serious weaknesses, both practically, because it may be very difficult to decide which of two crimes is the more serious and thus deserving of severer punishment, and morally, because if Deterrent and Reformatory considerations are altogether ignored when the list of penalties is drawn up a great social good might be sacrificed in order to achieve a small improvement in the accuracy of a punishment from the Retributive standpoint.   But, on the other hand, I have pointed out that the charge that Retributive theories of penalty-fixing are barbarous is based on the mistaken assumption that the only such theory is the *lex talionis*, and that a modified Retributive theory is perfectly possible, one which only uses Retributive considerations to fix some sort of upper limit to penalties and then looks to other factors to decide how much and what sort of pain shall be inflicted.   Purely Reformatory or Deterrent theories of penalty-fixing, which lack that limit, run the risk of becoming far more inhumane than even a purely Retributive theory.

Finally, I have been concerned to show that only if one subscribes to a Retributive theory of the moral justification of punishment (Problem 2) has one grounds on which to object to taking the Deterrent and Reformatory theories of penalty-fixing (Problem 3) to their logical conclusion, *i.e.* to inflicting pain on the innocent as a deterrent to others and as a means to removing suspected criminal tendencies before they can be manifested in actual offences. Those who object that such action would not (logically) be punishment are not objecting to the action taking place but only to its being given a certain name. Yet surely most people feel that there is more difference between inflicting pain on the guilty and inflicting pain on the innocent than that one can and the other cannot be called punishment?

University of Oxford

CK

e sub-
unish-
aking
fixing
in on
oving
ed in
d not
aking
surely
icting
that

IV.—OPPOSITE CONDITIONALS AND DEONTIC LOGIC

By P. B. Downing

I

IF one speaker utters the words ' If it rained the match would be cancelled ' and another speaker replies with the words ' If it rained the match would not be cancelled ', then provided that both speakers understand the language, and there is no misunderstanding as to which match is being referred to, it seems natural to conclude that two statements have been made which cannot, logically, both be true.   Indeed, it seems that the second speaker in such a conversation might equally well express himself by saying to the first speaker something like ' No, you are wrong, the match would not be cancelled ', and this surely suggests that what he intends to express contradicts the first speaker's statement. Yet in spite of such usages some philosophers hold the view that opposite subjunctive conditionals (by which I mean pairs of subjunctive conditionals of the form ' If p then q ' and ' If p then $\sim$q ', or of the form ' If p then q ' and ' If p then r ' where q and r are contraries) are logically compatible.   In the first half of this article I propose to resist some arguments in favour of this view and advance some reasons for holding the opposite view.   My remarks will be confined to conditional statements which are nontruthfunctional and ' empirical ' in that they assert, typically, that if some state of affairs, logically, but not necessarily physically possible, were to arise, continue, or come to an end, some other state of affairs would arise, continue, or come to an end.

Professor Popper [1] has attempted to establish the compatibility of opposite subjunctive conditionals by showing that they can be deduced from consistent sets of premises.   Of the two illustrative arguments which he gives I shall consider only the second, since only the second is concerned with empirical statements.   Popper asks us to consider the following four premises :

II   If Peter were now in his flat, he would be either in his bedroom or in his bathroom, and he could not be in both at the same time, and he could be in no other room (say, because the flat has no other rooms).

[1] MIND, lxviii (October 1959), 518.

    IIA If Peter were now about twenty miles from his flat, he
        could not possibly be in his bedroom now.

    IIB If Peter were now about twenty miles from his flat, he
        could not possibly be in his bathroom now.

    II+ Peter is now about twenty miles from his flat.

Popper thinks that if we consider, in the light of these premises,
whether or not Peter would be in his bathroom now if *per impos-
sibile* he were in his flat, it is possible to deduce from II, IIA and
II+ that if Peter were now in his flat, he would now be in his bath-
room, and from II, IIB and II+ that if Peter were now in his flat,
he could not possibly be in his bathroom, and that since the four
premises are consistent this shows that opposite subjunctive
conditionals are compatible.

    The four premises are complicated by certain logical connec-
tions. Thus seemingly II+ entails the statement that Peter is
not now in his flat, and each of the statements that Peter is now
in his bedroom and that Peter is now in his bathroom entails the
statement that Peter is now in his flat. Hence IIA and IIB are
necessarily true statements. Combined with II+ they appear to
yield no statements of physical impossibility, but, respectively,
the statement that Peter is not now in his bedroom, and the
statement that Peter is not now in his bathroom. This being
the case, it is hard to see why Popper thinks that a conditional
statement to the effect that if Peter were now in his flat he would
now be in his bathroom follows from II, IIA and II+. Premise
II and the statement that Peter is not now in his bedroom jointly
entail the disjunction that either Peter is not now in his flat or
he is now in his bathroom, but no conditional statement follows
from this disjunction. II, IIA and II+ jointly entail that Peter
is not now in his flat, and *a fortiori* neither in his bedroom nor in
his bathroom, but tell us nothing as to what would be the case if,
per impossibile, Peter were in his flat. For similar reasons no
conditional statement follows from II, IIB and II+.

    Popper appears to think that the four premises jointly entail
the impossibility of Peter's now being in his flat. If I am right
they entail the statement that Peter is not now in his flat, but
could only entail the impossibility of his being there in some very
weak sense such that the impossibility of p is deducible from its
falsity. However, it is easy to construct a set of premises which
resemble Popper's premises, but which obviously yield statements
of physical impossibility. Consider, for example, the following
statements :

(1) If Peter were in London by six o'clock, he would be either in his flat or at his club before seven o' clock.

(2) It is physically impossible that Peter should be in his flat before seven o'clock.

(3) It is physically impossible that Peter should be at his club before seven o'clock.

From (1) and (2) we can now deduce the statement that it is physically impossible that Peter should be in London by six o'clock and not be at his club before seven o'clock, from (1) and (3) we can (obviously) deduce the statement that it is physically impossible that Peter should be in London by six o'clock and be at his club before seven o'clock, and from (1), (2) and (3) it follows that it is physically impossible that Peter should be in London by six o'clock. However, the fact that the physical impossibility of Peter's being in London by six o'clock is deducible from these data does not show that any conditional statements are deducible from them. Conditional statements can take the form of speculations as to what would happen if something were to happen which is logically possible but physically impossible, and it is surely clear that although some statements of this kind may be true, others may be false, and that to be told that the antecedent of such a statement cannot be fulfilled is not to be told what would happen *if* it were fulfilled. Consequently from the fact that data (1), (2) and (3) yield a statement of impossibility it does not follow that they yield a conditional conclusion. Moreover, since ' p is physically impossible ' does not entail ' If p then q ', but *does* entail ' p.$\sim$ q is physically impossible ', where q is some proposition such that $\sim$ q is logically compatible with p, it follows that ' p.$\sim$q is physically impossible ' does not entail ' If p then q '. Hence, although (1) and (2) jointly entail the impossibility of Peter's being in London by six o'clock and not at his club before seven o'clock, it does not follow that they entail the conditional statement that if Peter were in London by six o'clock he would be at his club before seven o'clock, and similar objections can be raised against the deduction from (1) and (3) of the conditional statement that if Peter were in London by six o'clock he would not be at his club before seven o'clock. Indeed, more direct objections can be raised against these deductions. On any usual sense of ' physically impossible ', ' r is physically impossible ' does not entail ' If p then $\sim$ r ',though jointly with ' If p then r ' it appears to entail ' p is physically impossible '. Hence the deduction from (1) and (3) breaks down. Similarly, ' q is physically impossible ' and ' If p then q v r ' do not entail ' If p then r ', though jointly

with ' If p then $\sim$ r ' they appear to entail ' p is physically impossible '. Hence the deduction from (1) and (2) breaks down. It seems, indeed, that the only way to make these deductions work is to adopt an unusually strong interpretation of impossibility. For example, if we interpret (3) as something like ' Whatever the circumstances Peter would not be at his club before seven o'clock ' or ' For any proposition p, if p were true Peter would not be at his club before seven o'clock', we can obviously deduce that if Peter were in London by six o'clock he would not be at his club before seven o'clock. If (2) is interpreted similarly the course is not so clear. Admittedly we can deduce that if Peter were in London by six o'clock he would not be in his flat before seven o'clock, but our further deductions appear to involve the contrariety of opposite subjunctive conditionals, for it seems that if we cannot argue from the truth of the statement that if Peter were in London by six o'clock he would not be in his flat before seven o'clock to the falsity of its ' opposite ', we cannot, with the help of (1), prove the truth of the statement that if Peter were in London by six o'clock he would be at his club before seven o'clock. Be this as it may, the adoption of so strong an interpretation of impossibility renders (1), (2) and (3) inconsistent, since given that Peter could not be in his flat before seven o'clock in any conceivable circumstances and could not be at his club before seven o'clock in any conceivable circumstances, it cannot be true that if he were in London by six o'clock he would be either in his flat or at his club before seven o'clock, for it follows from this latter statement that either there are some conceivable circumstances given which Peter would be in his flat before seven o'clock, or there are some conceivable circumstances given which Peter would be at his club before seven o'clock. Hence, from (2) and (3) we are bound to deduce that (1) is false, and we have therefore not shown that opposite subjunctive conditionals can be deduced from consistent sets of premises.

The distinction between subjunctive conditionals and statements of physical impossibility is of general importance in the assessment of attempts to prove the compatibility of opposite subjunctive conditionals, since the plausibility of such attempts often depends on a confusion between the two kinds of statement. Statements of impossibility are sometimes expressed by means of subjunctive conditional sentences, and since a statement of the form ' p . $\sim$ q is physically impossible ' is compatible both with a statement of the form ' p . q is physically impossible ' and with a subjunctive conditional of the form ' If p then $\sim$ q ', this makes it possible to produce pairs of sentences which appear to express

opposite subjunctive conditionals which are compatible. Thus one might well be tempted to express the statements that it is impossible for Peter to be in London by six o'clock and not be at his club before seven o'clock, and that it is impossible for Peter to be in London by six o'clock and be at his club before seven o'clock by means of the sentences ' If Peter were in London by six o'clock he would be at his club before seven o'clock ' and ' If Peter were in London by six o'clock he would not be at his club before seven o'clock '. This mode of expression makes it look as if we have a pair of compatible opposite subjunctive conditionals, but if the distinction which I have been stressing is a valid one this appearance is deceptive.

My next task is to provide some positive support for the view that opposite subjunctive conditionals are contraries. I doubt whether a satisfactory proof of this view can be given, if by a satisfactory proof is meant a deduction from premises, such that both the premises and the deduction are more obviously acceptable than the conclusion, but I think that it is nevertheless possible to put forward good grounds for the view. This can perhaps be shown by an example. If we ask someone what he will do tomorrow afternoon, he cannot without contradicting himself reply that he will both go to the cinema and not go to the cinema. Suppose, now, someone insists that no such contradiction arises, on some extraordinary ground such as that contradiction is only possible in the case of necessary statements. It would be hard to refute him by means of a proof, since any premises used are likely themselves to be questioned by such a person. We should, therefore, have to adopt different tactics, and we might, for example, try to persuade him that not only, say, what numbers divide into thirty-six, but also what is going to happen tomorrow afternoon is something about which disputes are possible, and that to answer in the manner objected to is, as it were, to dispute with oneself. To say this is simply a persuasive way of stating, or hinting at, the conclusion which we are supporting, but it seems that it is only by persuasive methods that we can convince someone of the applicability to such cases of the principle of contradiction. The same is true of conditional statements. If we ask someone what he would do tomorrow afternoon if he were given the afternoon off, he cannot without contradicting himself reply that he would both go to the cinema and not go to the cinema. If someone denies this on the ground that opposite conditionals are compatible, a proof that he is wrong will be hard to give, but one can point out that it is possible to disagree not only about what *will* happen tomorrow afternoon, but also, and in just the same way, about

what *would* happen tomorrow afternoon if some given thing were
to happen.   To understand that disputable territory includes
' what would happen if ' or ' what would be the case if ' is part of
what is required in order to understand in detail how the principle
of contradiction applies to conditional statements, but if someone
refuses to understand this, it seems that we can only use per-
suasive arguments against him.   However, if, as seems highly
likely, our interlocutor had had some logical training, we could
perhaps back up our persuasive arguments by pointing out that
the view that opposite conditionals are not contraries arises from
the confusion of conditional statements with disjunctions dressed
up as conditional statements, or with statements of impossibility.

I conclude that good reasons can be put forward for the view
that opposite subjunctive conditionals are contraries, and that to
ask for a proof in some further sense is as inappropriate as to ask
for a proof of other applications of the principle of contradiction.
If the contrariety of opposite conditionals is a logical truth, it can
of course be used as a basis for further deductions.

II

It seems that opposite conditionals which are contraries also
occur in moral discourse, and that this view can be defended along
lines similar to the above.   If we ask someone whether it is our
duty to help our neighbours, he cannot without contradicting
himself reply that it is both our duty to help our neighbours and
our duty not to help our neighbours.   Similarly, if we ask some-
one whether if we harmed a neighbour it would be our duty to help
him, he cannot without contradicting himself reply that it would
be our duty both to help him and not to help him.   Of course, the
contrariety of opposite categorical moral judgments of the
form ' It is obligatory to do A ' and ' It is obligatory not to do A '
is disputed by subjectivists, but although this view can be main-
tained without inconsistency, it leads its protagonists into giving a
very unconvincing account of moral disagreements.   I shall not
now take up this theme, but shall assume, without argument, that
opposite categorical moral judgments are contraries.   My point
then is that if we accept this view, then if we understand the
nature of conditionality we cannot honestly avoid taking the
further step of admitting that opposite conditional moral
judgments are also contraries.

This has a bearing on the analysis of moral judgments as im-
peratives.   As Mr. Hare has suggested, there seems to be no good
reason against saying that pairs of commands are contraries or

contradictories. Hence, one advantage of the imperativist analysis is that it can preserve the contrariety of opposite categorical moral judgments, since these can be held to be equivalent to commands of the form ' Do A ' and ' Don't do A ' which are contradictories, or of the form ' Do A and B ' and ' Don't do A ' which are contraries, and so on. Difficulties emerge on further investigation, however. The logical relations of commands depend on the logical possibility or impossibility of complying with them. Thus, for example, 'Give me a good book' entails 'Give me a book ' because it is logically impossible to comply with the first without complying with the second, and 'Give me a book' and 'Don't give me a book ' are contradictories because it is logically impossible either to comply with both or to comply with neither. Sometimes, indeed, it is suggested that the contradictories of commands are permissions, but this view seems to result from confusing commands with statements about commands, or the contradictories of commands with cancellations or withdrawals of commands. Thus the contradictory of ' A is commanded ' interpreted as a statement, is no doubt ' $\sim$ A is permitted ', but the contradictory of ' Do A ' is nonetheless the command ' Don't do A ' Moreover, to cancel a command is not to issue its contradictory any more than to withdraw a statement is to make its contradictory statement. If the principle that the logical relations of commands depend on the logical possibility or impossibility of complying with them is applied to so-called conditional commands it becomes clear that here the logic of commands and the logic of moral judgments part company, since opposite conditional commands are not contraries. Thus ' If you come home late lock the door ' and ' If you come home late don't lock the door ' are compatible, since it is logically possible to comply with both by not coming home late. The fact that this may for some reason be physically impossible is, of course, irrelevant. Indeed, it seems that ' If you come home late lock the door ' can be interpreted as ' Either don't come home late or lock the door ' or as ' Don't both come home late and not lock the door ', and if the dubious expedient is adopted of regarding as a command a disjunction of a command with a statement, so that the conditional command becomes ' Either you will not come home late, or lock the door ', the two opposite conditional commands will still be compatible, since it will still be logically possible to disobey neither. It seems, therefore, that the only conditionality applicable to commands is a spurious conditionality akin to that of material implications. It is, of course, possible to interpret conditional commands as commands issued conditionally, that is, we can say

that certain sentences only operate to give a command if certain conditions are fulfilled. Thus the sentence 'If you come home late lock the door ' can be interpreted in such a way that if the person in question comes home late a command to lock the door comes into operation. In such a case, however, the only command involved is a categorical command, and indeed we obviously cannot, by laying restrictions on the use of a sentence, invent some special kind of conditional command. It may be added that an interpretation of conditional moral judgments in terms of conditionally issued commands would be quite unacceptable, since if we utter some such words as ' If anyone made a promise it would be his moral duty to keep it ' we are clearly making a moral judgment the making of which does not depend on the fact that promises have been or will be made.

It seems, therefore, that the incompatibility of opposite conditional moral judgments provides a reason for rejecting the view that moral judgments are commands, and perhaps for rejecting the view that they are in any important respect like commands. Moreover, although my arguments for the contrariety of opposite conditionals do not provide a proof of their contrariety, they perhaps provide a kind of explanation of it. If they are accepted as such they suggest strongly that moral judgments are statements of fact, for the incompatibility in question is ' explained ' by saying that disputes are possible not only about what is the case, but also about ' what would be the case if ', and it seems that the notion of what would be *the case* is essential to any such explanation.

The character of opposite conditionals is also relevant to some recent discussions of Professor von Wright's system of deontic logic.[1]   The sytem in its original form made no allowance for conditional moral judgments, but after Professor Prior had pointed out that it gave rise to deontic paradoxes similar to the paradoxes of material implication, von Wright amended the system by introducing the notion of permission or obligation under certain conditions, and he suggested that an axiom might be introduced to the effect that under any given conditions, one is either permitted to do, or permitted to neglect any given act.

It might be thought that this axiom does justice to the incompatibilities which I have been stressing, since it rules out, for

[1] See von Wright's article in MIND, lx (January 1951), 1 (reprinted in *Logical Studies*). Also discussions by Prior and McLaughlin in MIND, lxiii (January 1954), 64 and lxiv (July 1955), 400, and von Wright's reply in lxv (October 1956), 507.

example, the possibility that someone's harming his neighbour might commit him both to helping and to not helping his neighbour. However, an examination of von Wright's original article and subsequent comments suggests that the introduction of the axiom is only a palliative remedy. The original article gives the impression that the system is conceived not as a system of commands, but as a system of statements about commands. Von Wright uses words such as ' permitted ' and ' forbidden ', and the contradictory of OA (A is obligatory) is not O $\sim$ A ($\sim$ A is obligatory), but P $\sim$ A ($\sim$ A is permitted). If one were considering the logic of statements about the commands given by some authority, these are the sort of features one would expect to find. Moreover, it seems that von Wright's Principle of Deontic Distribution would be applicable to such a system, since it is reasonable to assume, for example, that if the authority issues a command, it issues all commands entailed by that command. However, in such a logic we should not be entitled to assume that ' A is commanded ' and ' $\sim$ A is commanded ' are contraries, for the simple reason that to make statements from which it follows that the authority is inconsistent is not to fall into inconsistency oneself. Moreover, if we stipulate that the authority is consistent, this stipulation is not a logical truth but contingent in character, for it is clearly a contingent question whether or not some authority makes consistent pronouncements. When von Wright comes to consider this problem, he writes rather as though he were hesitating between this kind of logic and something else. He asks whether it is ' illogical ' to accept the consequences of the view that an act and its negation can both be forbidden, and says that ordinary language and our common sense logical intuitions seem at first not to help us. He then points out that acceptance of the view involves the consequence that an act can be both obligatory and forbidden, suggests that an appeal to ordinary usage will lead us to reject this, and introduces a principle, The Principle of Permission, to the effect that any given act is either itself permitted or its negation is permitted. Later he says that in the case of some of the laws of deontic logic, it is not intuitively clear whether their truth is a matter of logic or moral code, and it seems that here he must be thinking of laws which depend on the Principle of Permission.

Yet surely no laws of deontic logic, whether or not connected with a Principle of Permission, can be a matter of ' moral code '. If I am right, OA and O $\sim$ A interpreted as statements about commands are not contraries, and a contingent principle has to be introduced to ensure their incompatibility, but if they are

interpreted as statements of moral obligation they are contraries, this being a matter not of moral code but of logic. Von Wright's hesitation over this problem suggests the possibility that he is thinking in terms of statements about commands, and hence suspects that the Principle of Permission is not a logical truth, but that he adds it in order to bring his system closer to the logic of moral obligation. Such an expedient would, of course, be inadequate, for to say that opposite categorical moral judgments cannot both be true given the truth of some contingent statement, is not to say that they are contraries. Moreover, on this view of the matter, the new axiom must be regarded with some suspicion. It seems that if the notion of permission or obligation under certain conditions were introduced into the logic of statements about commands it would have to be understood in terms of the conditionally issued commands which I have already discussed, and that the new axiom would be a contingent statement guaranteeing the consistency of the authority in its conditionally issued commands, just as the Principle of Permission guarantees its consistency in its unconditionally issued commands. Hence, for reasons similar to those given in the case of the Principle of Permission, the new axiom could not do justice to the contrariety of opposite conditional moral judgments.

But is von Wright really confusing the logic of moral judgments with that of statements about commands ? Perhaps not, but to suggest that he is is much the same as suggesting that he confuses moral with legal obligation, and it can at least be said that his system is much more like the logic of legal obligation than like the logic of moral obligation. For one thing, the system could easily be interpreted in terms of statements about some set of laws assumed to be consistent. For another, the logic of moral obligation, if I am right, is concerned with what would or might be the case, whereas statements about the law ' as it stands ' are only concerned with what *is* the case, and that von Wright is only concerned with what *is* the case appears to be borne out by his reactions to an objection brought against his system by Mr. McLaughlin.[1] McLaughlin queries the law OA . O (A ⊃ B) ⊃ OB on the ground that from OA and O (A ⊃ B) it does not follow that B is obligatory even though A is not performed. Von Wright replies that it *does* follow, but is this true ? Let us first consider the validity of the law itself. It can certainly be interpreted as a truth of the logic of legal obligation, and can, indeed, be proved without the aid of contingent assumptions, but can it be interpreted as a law of the logic of moral obligation ? As soon as one

[1] MIND, *loc. cit.*

thinks of moral obligation in terms of ' what would be the case if ' it becomes possible and appropriate to accept the view that to say that an action A is absolutely obligatory is to say that it would be anyone's duty to perform it whatever the attendant circumstances, and that to say that an action B is absolutely obligatory under conditions A is to say that if A occurred it would be anyone's duty, whatever the attendant circumstances, to perform B.   On this view O(A . B) does not entail OB, since the former, but not the latter, is consistent with O (If ∼ A then ∼ B).   It may be absolutely obligatory to perform A and B, but also absolutely obligatory, under conditions ∼A, not to perform B, in that if anyone neglected to perform A it would be his duty, whatever the attendant circumstances, not to perform B.   But if B is absolutely obligatory there cannot be *any* circumstances in which it would be anyone's duty not to perform B.   If OA, OB and O(A . B) are given this meaning, and if O(A ⊃ B) is interpreted as something like ' Any circumstances in which A is absolutely obligatory are circumstances in which it is absolutely obligatory, if possible, also to perform B ', then seemingly OA . O(A ⊃ B) entails O(A . B). But there seems to be no plausible way of interpreting O(A ⊃ B) which will make OA. O(A ⊃ B) entail O(If ∼ A then B), and since OB entails O(If ∼ A then B) this means that OA . O(A ⊃ B) cannot entail OB.   Hence, on this interpretation of absolute obligation, McLaughlin would be right to dismiss the law as invalid, and right to give as a reason for this the fact that it does not follow from the premises that B is obligatory even though A is not performed. It is, of course, possible to interpret the law in such a way that it is valid, by interpreting OA as ' a set of actions including A is absolutely obligatory ', and OB correspondingly, but even if this is done, McLaughlin's contention that O(If ∼ A then B) does not follow from the premises is correct.   Indeed, as soon as the possibility of genuinely conditional moral judgments is allowed for, it becomes clear that McLaughlin is right on *this* point, since O(A . B) is consistent with O(If ∼ A then ∼ B).   To return to the interpretation of OA and OB as statements of absolute obligation, it seems that a more radical objection than McLaughlin's can be raised against the law, for if A and B are logically independent, it seems that OA and OB are incompatible, since it might be that someone is placed in circumstances in which the only way to perform A would prevent the performance of B.   In such circumstances, given OA, it would be his duty to perform A.   Hence, given OA, it could not be his duty to perform B in any circumstances whatever.   Thus it can be argued that what follows from OA . O(A ⊃ B) is not OB but its negation.

If these interpretations are justified, the logic of absolute obligation turns out to have some curious features. Nevertheless, this kind of logic is not so far removed from our everyday reasonings as might be supposed. If we are told that if A happened then *inter alia* B would happen, this information may be very useful, especially if we are primarily interested in ensuring the occurrence of B. On the other hand, if we are told that if we performed A it would be our duty, *inter alia*, to perform B, this is somewhat unhelpful. If we make no further inquiries, but perform both A and B, we may well infringe some moral rule, for it may be that we have neglected to perform some other action C, and it may be that although if we performed A it would be our duty to perform C and B, and hence *inter alia* to perform B, nevertheless if for any reason we performed A but failed to perform C it would be our duty not to perform B. In such circumstances it seems that however many statements of the *inter alia* type are put forward, similar doubts may remain, and that the only way to provide a final answer is to put forward either a statement of absolute obligation to the effect, perhaps, that if we performed A it would be our duty, whatever the attendant circumstances, to perform B (such a statement, though conditional in form, is ' absolute ' in its insistence that nothing is to be taken into account other than what is specified), or possibly a statement concerning both absolute obligation and causal or near-causal connections, such as the statement that if we performed A circumstances would arise given which it would be our duty, whatever other circumstances arose, to perform B.

These last reflections go beyond the contrariety of opposite conditionals, and raise problems into which I do not now propose to enter. I suggest, however, that they provide further support for the view that the logic of moral obligation differs radically both from the logic of legal obligation and from the logic of imperatives.

University College, London

V.—UNEXPECTED EXAMINATIONS AND UNPROVABLE STATEMENTS

By G. C. Nerlich

There have been several articles in Mind recently on the paradox of the unexpected examination. Despite the fact that this paradox appears a thoroughly trifling affair it has, unexpectedly, withstood examination. My view is that it is by no means either trivial or easy to solve. Moreover, I think it is a quite unique kind of ordinary language problem, having some connection with the situation posed by Goedel's famous sentence, to the effect that the sentence itself cannot be proved. It will be clear, when I have dealt with the paradox, why I think it is of some importance to logic—of more importance than the comparatively simple Grelling paradox, for example.

In the first part of what follows I devote some space to what has already been written about the paradox, so as to clear away what is inessential, and also to make more acceptable my own rather bizarre suggestion as to the solution, which is given in part II. In the third part there is an attempt to present the argument in a wholly formal way. The last section concerns a special problem arising out of my solution, which is that an unexpected examination can be held on the last day. Why this deserves a section to itself will soon be clear.

I

The paradoxical state of affairs is the following : A headmaster announces to his school that an unexpected examination has been arranged for one of the five school days of the ensuing week. It will be unexpected in the sense that the pupils will not know when it is to occur prior to its occurrence. The pupils accuse the head of contradicting himself, and justify the charge in the following manner : No *unexpected* examination can occur on Friday, the last day of the week, for on the Thursday evening the scholars would know the examination must occur on the only remaining day, so they would be expecting it on Friday. But, having reached this conclusion, it is now clear that no *unexpected* examination can occur on Thursday either, for on Wednesday evening the pupils, having seen that Friday is eliminated could now see that it must be arranged for the following day, so the Thursday examination would not be unexpected. Hence Thursday and Friday are both

503

out. By repeating this pattern of argument they succeed in showing that no day is possible for an *unexpected* examination, so that the head has contradicted himself.

Yet it is quite clear that the head has *not* contradicted himself. Suppose he has arranged the examination for Wednesday. The pupils have no way whatever of finding out that this is the chosen day. Nothing points to Wednesday rather than, say, Tuesday or Thursday as the day on which to have their pencils and wits especially sharp. Something has gone wrong in the argument. But what ?

If we attend to the reasoning it is obvious that everything of interest centres round the word " unexpected ", and it is a primary task to get it clear. Now the sense in which " unexpected " is actually used here is " it cannot be *deduced* from the announcement when the examination has been arranged ". For the pupils appear to falsify the head's statement by arguing that, prior to any day D, it can be *deduced* from the statement that the examination will be held on the day in question. So that the head makes a claim that the examination date does not follow from his statement, and the pupils use *just that claim* to " prove " that it does follow. Accordingly, I shall adopt the following gloss on the statement we are concerned with :

(1) An examination has been arranged for one of the five days of week W.

(2) It is not possible to deduce from (1) and (2) on which day the examination has been arranged at any time prior to its occurrence.

A very noticeable feature of this as it now stands, is that (2), the second member of the conjunction, refers not only to the first member (1), but also to itself. Which is an oddity.

Just this was thought by Mr. Shaw in a most illuminating treatment of the puzzle, to be the source of the paradox. He said that (2) can only properly and consistently state that the date cannot be deduced *from* (1) (MIND, July 1958). Shaw considers the situation as embodied in a set of school rules. Rule 1 decrees an examination in week W. Now if the second rule, which is about unexpectedness, decrees only that the examination must be arranged so that it cannot be deduced from Rule 1 when the examination is to take place, then no contradiction follows, and so no paradox results. Under these two rules one cannot hold an unexpected examination on Friday, but any other day will do. For the contradiction to follow, we need to use Rule 2 as a premiss

from which to deduce the date, but we can only do so if it refers to itself. Shaw seems to take it for granted that no statement may refer to itself, though Goedel would seem to have demonstrated that at least one can.

In a way this is beside the point, though interesting enough in itself, since the remainder of Shaw's argument demonstrates that the self-referential feature is quite inessential to the paradox. He seems not to have seen this. It is shown quite clearly how to get the paradox afoot without employing a statement which refers to itself at all. It is done thus. To Rule 1 and Rule 2 (which refers only to Rule 1) we add four more school rules, Rules 3-6. They are all of a similar kind to Rule 2. Rule 3 says that one can't deduce the date from Rules 1 and 2. Each rule says that the date cannot be deduced from any previous rules, so that Rule 6 states, in full, that the examination will occur on a day such that it is not possible to deduce from Rules 1-5 on which day it will occur at any time prior to its occurrence. (The point is to have the first rule, and as many further rules as there are days.) From this set of rules a contradiction follows by a precisely parallel series of steps as are used to derive a contradiction from my statements (1) and (2). Shaw claims that there is no *paradox* here since no unexpected examination can *in fact* be set in week W. But, in fact, one *can* be set. Certainly it is possible for an examination to *occur* during the week, say on Wednesday. And it is quite plain that there is no way of deducing from Rules 1-6 that Wednesday will be the day. So an examination *will* be unexpected if it occurs then. The pupils can deduce neither more nor less from Rules 1-6 than they can deduce from my (1) and (2). So this state of affairs is paradoxical—a contradiction does follow from plainly consistent premises—yet no self-referring statement is involved. I shall stick to the original formulation, *i.e.* to (1) and (2), as I think that is the most natural and proper interpretation of the head's statement. But, as I am at pains to point out here, nothing essential to the formulation of the paradox is entailed by the curious self-referring nature of (2).

Before I move on to the next point of criticism, let me remark that the problem is already becoming a little clearer. It seems that by adding statements to the effect that something is not provable, we thereby make it provable. But, in fact, we add no information to the purpose in adding such remarks. For we still fail to be able to deduce, say, a Wednesday examination. But it looks fairly plain at this stage that *provability* is our bugbear.

Mr. Lyon argues that the pupils' reasoning contains a fallacy of equivocation. In his view, the statement that the examination

will be unexpected " entails, or if you like, means "[1] roughly one of the following :

(a) it will not be possible to deduce from the statement when the examination will occur at any time prior to its occurrence, *unless it occurs on the last day.*

(b) it will not . . . , *whether or not it occurs on the last day.*

Lyon claims for obvious reasons, that (b) is inconsistent with the statement. But if the head means (a), which Lyon regards as the only consistent thing he can mean, then the paradoxical argument fails. For the pupils overlook the " unless " clause which permits the head to hold the examination on the last day even though the students can deduce on the previous evening that it will be held then.

But this is surely not permissible. Whether or not the head's statement entails (a), it certainly does not *mean, i.e. is not equivalent to* (a). For there is a strict and perfectly proper sense of " unexpected " in which (a) is equivalent rather to " the examination will occur unexpectedly, unless it occurs expectedly on the last day ". Whereas the head can plainly mean strictly what he says—that the examination will occur unexpectedly. In denying that (a) is what the head says, one is not, of course, asserting that he says (b). For one is not denying, *ipso facto,* that if the examination *were* to occur on the last day then it would be possible to deduce the fact.[2] One is denying only that it will occur on *any* day prior to which such a deduction is possible, which seems to mean denying that it will occur on the last day. And now we are off again on the paradoxical argument. Yet, let me insist again, the head does not contradict himself. For if he has arranged the examination for Wednesday (which would make the " unless " clause of (a) simply otiose and trivialising), the pupils have no way at all of deducing the date of the examination in advance of its occurrence. So that the examination would *not* occur on any day prior to which the deduction is possible, which is what the head stated. He does not, therefore, need to leave open the possibility that he will set the examination expectedly in order to set it unexpectedly. To drive this point firmly home, there are tests which actually *require* the rejection of the " unless " clause and such tests occur daily. The trial emergency stop in every driving test is a case in point. The trial is improper if the order does not take the candidate unawares, so it cannot be

---

[1] Ardon Lyon, MIND, October 1959.
[2] But see section IV of the present paper.

allowed to occur expectedly even at the end of the test. Yet proposing such a trial is not proposing anything contradictory.

My last critical point is directed against a possible suggestion that the students cannot rule Friday out as a day for an unexpected examination until Thursday evening is actually reached. One cannot, it might be urged, argue in this way at the time of the announcement of the examination. Only *if* they reach Thursday without the examination can they argue in this way. But it can readily be shown that this condition is irrelevant to ruling out a Friday examination; for example, by the following dilemma. Either the examination has been held prior to Thursday, or it has not. If it has, then the examination is over and will not take place on Friday, *i.e.* no examination expected or unexpected will occur on Friday. If it has not been held prior to Thursday evening, then it follows from the head's remark that it will be held on the morrow, so that no unexpected examination will occur then. (This assumes that the head has not been playing cat-and-mouse with his pupils and does not mean to set two exams. So diabolical a twist need not concern us.)

II

I turn now to my own view of the solution to our problem. Briefly and a little loosely, it is this. The pupils' argument consists of stages, and at the end of each stage they get the *negation* of a statement of the form " Examination on –day ". So that all they can deduce in the argument is that none of the days is a day for which the examination can have been arranged. But then, if they take it as proved that no unexpected examination can be set, any examination which *is* set will find them unprepared, and so it will be unexpected. That is, any day excluded in the argument satisfies this description : " a day such that it is not possible to deduce from the head's statement, at any time prior to the day, that the examination *has* been arranged for this day ". Since all we get in the argument are statements which are *negations* of dates for the examination, we never get one of the right form to falsify the head's statement. And once we get to the position where we have deduced the negation of each of the alternates set out in (1) of the remark, it is obvious that there is no further deduction to be carried on. We have gone as far as we can go, and are still without the examination date. (I ignore the possibility of every statement's following from a contradiction. This is of no use to the pupils.) So *due to the fact* that it entails not, *e.g.* Examination on Wednesday, but something else (a contradiction), the statement is self-consistent.

This is a hard saying. However, let us look again at the curious logical features of this everyday remark. The statement is partly about an examination and partly about its own logical consequences, *viz.* that the examination date is not among them. So that part of it is metalogical to itself as a whole. What (2) says, metalogically, is that one of the five statements included in (1) is true but not provable using either (1) or (2) as premisses in an argument. The only way in which this metalogical statement can be falsified is by proving that the examination *has* been arranged for a certain day. It is this that the students attempt to do but fail to do, producing only days on which it seems *not* possible to hold it. And that is because in the attempt, they are forced to use the very premiss (or set of premisses) which they hope to falsify. I do not mean that this alone can be what produces our queer state of affairs, since it is standard procedure in *reductio ad absurdum* arguments and is plainly valid. What does produce the oddity is that the premiss used states its inability to be used in that way ; for it says that if any statement of the five can be deduced then it is false, *i.e.* if p can be proved so can not-p. So that to embark on this preposterous logical voyage is bound to land us in trouble. (Alternatively, if we have no self-referring premiss but an ordered set of premisses, then we find ourselves using premisses in the argument which later premisses state cannot be so used. And we must use the earlier premisses in the forbidden way *together with* the later ones if we are to get our contradiction.)

What is so unusual about this situation is firstly that it seems as if the metalogical premiss *can* be used in just the way the students require and the head denies ; whereas it would have been less interesting and more common if the metalogical premiss had neither yielded nor seemed to yield any step. Secondly it so turns out that the contradiction, which really is formally derivable, does not entail the contradictoriness of the conjunction from which it is derived. For if the examination is held on Wednesday (1) will certainly be true, and it is equally certain that (2) will be true, for where is the proof that the examination will occur then ? The statements are plainly consistent. What the contradiction does show is the futility of attempting to deduce the date of the examination in defiance of the metalogical premiss which has to be used in the course of the argument—to put the point informally.

In so far as it is about provability, this is a little like the state of logical affairs in Goedel's theorem. (Not in so far as it involves self-reference, since that is not central to our problem at all.) Goedel. proves that the sentence G is a proper sentence. But

what G says is that G is not provable, *i.e.*, it says this of itself. Goedel says that if this logical system is to be consistent then G must be undecidable, for if we prove G, then that proves that G is provable, which is the negation of G. The other side of the coin is that if we prove the negation of G, what we prove is that G is provable, *i.e.* G is not provable is provable, a contradiction again. So that if the system is consistent then it cannot be complete too, for if either G or its negation is provable, we get a contradiction. Our system is rather like this too. For we may look at (2) as saying that the system is not complete, *i.e.* that one of the five alternates of (1) is not provable but nevertheless is true. But just that remark about incompleteness seems to make the system now complete, and therefore contradictory. Yet, as we have seen, it is really neither complete nor inconsistent.

III

I claimed earlier that the contradiction really is formally derivable and I now want to go on to make the claim good by presenting the derivation in a rigorous way.

One of the striking things about the paradox, on the face of it, is that it has to do with time. The argument depends on the days' being temporally ordered. It is this fact which prompts the view I criticised earlier that no argument can be given to rule Friday out until Thursday evening is actually reached. I am inclined to think that the ordering of the *days* has nothing essential to do with the paradox, since it is only a way of bringing about an ordering of the *statements* " Examination on Wednesday ", etc. in respect of information correlated with them. What I mean is this : the statement that the examination is to be held on Wednesday has correlated with it the statements that it has not been held on Tuesday nor on Monday, *i.e.* the pupils know these statements to be true prior to Wednesday. So that if the examination has been arranged for that day the two negative statements are available for use in the deduction. I shall call this pair of statements the *deduction correlative* of the statement " Examination on Wednesday ". Each of the statements in (1) is, therefore, informationally ordered by way of the logical powers of its deduction correlative. All this is quite apart from the fact that the truth of any one of the statements entails the falsity of the rest, there being only one examination ; for that fact is not available for use in the deduction. In the present interpretation of the formal system I am about to set up, this ordering by way of information is achieved, in turn, by the temporal ordering of the days.

We are now in a position to drop the question of days and consider only statements in the course of the argument. Let the letters p, q, r, etc., represent respectively, " Examination on Monday ", " Examination on Tuesday " and so on. Then formally, (1) becomes :

(1) p v q v r v s v t.

The metalogical statement (2) becomes :

(2) $(\pi)\,[\{(1)\,.\,(2)\,.\,(dc\pi) \supset \pi\} \supset \sim\pi]$.

The variable π ranges over the statements of (1). $(dc\pi)$ means the deduction correlative of π. Premisses are assumed and dropped in the usual way by conditional proof.

I shall now give the argument preceded by the list of deduction correlatives.

$$dcp \equiv p\,v\sim p \qquad\qquad dcq \equiv \sim p$$
$$dcr \equiv \sim(p\,v\,q) \qquad\qquad dcs \equiv \sim(p\,v\,q\,v\,r)$$
$$dct \equiv \sim(p\,v\,q\,v\,r\,v\,s)$$

(1)	(1) p v q v r v s v t.	(1)
(2)	(2) $(\pi)\,[\{(1)\,.\,(2)\,.\,(dc\pi) \supset \pi\} \supset \sim\pi]$.	(2)
(3)	(3) dct.	(3)
(3)	(4) $\sim$(p v q v r v s).	(3)
(1) (3)	(5) t.	(1) (4)
	(6) (1) . (2) . (dct) $\supset$ t.	(1) (4) (5)
(2)	(7) $\{(1)\,.\,(2)\,.\,(dct) \supset t\} \supset \sim t$.	(2)
(2)	(8) $\sim$t.	(6) (7)
(9)	(9) dcs.	(9)
(9)	(10) $\sim$(p v q v r).	(9)
(1) (2) (9)	(11) s.	(1) (8) (10)
	(12) (1) . (2) . (dcs) $\supset$ s.	(1) (8) (10) (11)
(2)	(13) $\{(1)\,.\,(2)\,.\,(dcs) \supset s\} \supset \sim s$.	(2)
(2)	(14) $\sim$ s.	

Consider also this set of premisses :

(1) p v q v r.

(2) $(\pi)\,[\{(1)\,.\,(dc\pi) \supset \pi\} \supset \sim\pi]$.

(3) $(\pi)\,[\{(1)\,.\,(2)\,.\,(dc\pi) \supset \pi\} \supset \sim\pi]$.

(4) $(\pi)\,[\{(1)\,.\,(2)\,.\,(3)\,.\,(dc\pi) \supset \pi\} \supset \sim\pi]$.

Here no premiss refers to itself, but the contradiction is derivable in just the above way. Yet both these sets of premisses are consistent by interpretation.

I have not carried the argument to its end, since the pattern needs only to be repeated in a perfectly obvious way. But notice how odd the situation already is at line (14). Suppose it is

Wednesday evening and no examination has yet been set. The pupils have two alternatives open, a Thursday or a Friday examination and they cannot choose between them before they begin their argument. They now attempt to choose, arguing that Friday is not a possible day. But this attempt to choose puts them in precisely the same position *vis-a-vis* Thursday which they envisaged *vis-a-vis* Friday in order to rule it out. Making out a cast iron case against Friday results in making out the very same cast iron case against Thursday. And now they don't know what to think. It seems that *both* days are out but while it seems to be that way, then there is still nothing to choose between these days supposing that the head merely smiles at their reasoning and repeats his intention to set an unexpected examination. In short the argument yields nothing at all. It might as well not have begun.

Instead of saying what I said before, *viz.* that the contradiction is validly derived but is not a *reductio ad absurdum* of the premisses, we might choose to say instead that the argument is invalid, since to demonstrate the negation of any of the statements in (1) is thereby to demonstrate its possibility, on a higher level one might say, owing to the queer sort of statement (2) is. But the choice as to which we say amounts to little enough; it is no more than a recommendation in a new state of affairs. We haven't really shown that not-p doesn't follow, only that its following leaves us with possibly-p demonstrated on a higher level. So that we have not really avoided the contradiction, but merely thought of a way of making our encounter with it sound less objectionable. To sum this up: We might say that we have here a new form of reductive argument, for the valid derivation of the contradiction shows, on a higher level, that the *argument* stultifies itself. Or we may say that the argument is not valid, for deriving not-p is a higher level derivation of possibly-p. The choice here would depend on what one wanted to go on to do about it later, *i.e.* on what formal devices one would invoke. But, one might say too, the choice is really quite empty.

IV

Now what I have said creates a special problem about the last day. For it seems to leave open the possibility of an examination on that day in the following way. Suppose it is Thursday evening. We can deduce from (1) that the examination will be held next day. But (2) claims that there is no examination on a day such that it can be deduced that it will be held then. (1) and (2)

clash, so that it looks as if one of them must be false. But which of them is false ? Will there be an examination tomorrow only an expected one, or will there be no examination, because it could only be expected ? If we do not know which of (1) or (2) is false then we cannot be sure what the morrow holds. So that if what the morrow holds is an *examination*, that is something we cannot now prove. Therefore an examination will be unexpected. In that case (1) and (2) would both be true after all. But, we may now say, if there can be an unexpected examination tomorrow, then that is surely what is going to happen. (Here we jump to a new level.) However, we are expecting it again if we make this move, so there still seems to be a conflict between (1) and (2). We are back where we began, or at any rate, we are in a precisely parallel position in a higher sphere. We can go over the same ground again, but we will clearly strike the same situation on the next level. We seem to be able to decide nothing, not even that we can decide nothing.

All this makes one's head spin. Surely, one wants to say, it is mere verbal jugglery. But it seems plain to me that to allow Thursday as a possible day for an unexpected examination, one *must* take up this very uncomfortable position that Friday is also possible. One can avoid the difficulty over Friday only by encountering it over Thursday or elsewhere, though the shift may make it seem a less brutally stark encounter. But that would be only seeming. And it must be noticed that all along, the pupils have taken it as utterly certain that the examination *cannot* occur on Friday. What I have pointed out here is the role played by this certainty (this negative expectation) if the head does hold the examination on the Friday after all.[1]

This is not to deny that Friday is a queer case. For only premiss (1) is used on Thursday evening to deduce that the examination will occur next day. Someone might object, in conformity with the way I put the matter earlier, that the fallacy only arises when (2) is employed. He might go on, a Friday examination would be an expected examination but a Thursday examination would not, since, in order to be in a position where one can apparently show that Thursday is the day, one must use (2) to rule out

---

[1] At this point I wish to raise a question which I am not competent to settle. The situation for a *Friday* examination is quite closely parallel to the situation which would be posed by deducing G in Goedel's system. If what I have argued here is right, then this would not show the system to be inconsistent. If G were proved our next step would be ∼ G, a third would be G again and then, once more ∼ G, and so on. To prove G is also to prove it false. Which is not to prove it at all. The contradiction verifies that G cannot be proved.

Friday. But (2) states that it provides no such information as it is here being pressed to provide.

But this is to drop back into the pit. For the objector is now committed to the view that Thursday is the last possible day for the examination. The point is that we are saddled with (2) and cannot discount it on Thursday evening as part of what is given, except by a quite arbitrary prohibition. It would be no more arbitrary to prohibit the use of (1). Nor can we say that it is clear that Friday is not open to the head, but no-one is allowed to make use of the fact in an argument. Nothing justifies such a prohibition. The argument is not *forbidden*, but mistaken. For it results in deductive stalemate, the stalemate operating because we are saddled with (2), and it operates no less on Thursday evening than on any other. Which means that one can *not* deduce a Friday examination in advance.

From an informal point of view, however, there are circumstances under which the discounting of (2) might not be arbitrary, *e.g.* if the sources of (1) and (2) were different. If the head said (1), and (2) were added by an irresponsible junior master ; or again if the junior master said (1), and (2) were added by the head; we should feel in the one case entitled to expect an examination despite the addition of (2) and in the other we should not know at all what to expect. Again, to adapt a case of Mr. Lyon's, I may know (because I have seen) which cards are held in a player's hand. If he says that he will play a certain one unexpectedly, and leave that card till last then I have a right to discount his claim to play it unexpectedly. Though it may be worth noticing here, that since he appears to have contradicted himself, I have no kind of assurance that he will play the card at all. For *he* said he would play it, not anyone else, and his assertions no longer seem to be comprehensible. In the case we have been considering all along, one person is the source of all we know, and when our only source of information seems to impeach itself, then we do not know what to make of it. That is precisely what this particular source of our knowledge wants to achieve. One way of saying nothing is to contradict yourself. And if you manage to contradict yourself by saying that you are saying nothing, then you do not, in the end, contradict yourself at all. You *can* eat your cake, and have it too.

University of Leicester

VI.—ON NECESSARY TRUTH

By D. W. Hamlyn

WHAT is it to say that some truth is a necessary truth? In a recent article [1] Mr. Richard Robinson raises a number of difficulties about the notion of ' necessary propositions '—pointing out that the notion has been used historically in a number of different ways ; that Kant's formulation of the notion in his account of what it is for a proposition to be analytic is both obscure and muddled ; and that the Positivist distinction between empirical and analytic propositions rests upon a principle of division which is in fact double. He proposes to retain only one of the analytic/synthetic and necessary/contingent distinctions, and while finding defects in each he prefers the distinction originally instituted by Leibniz between necessary truths whose opposite implies a contradiction and contingent truths whose opposite does not.

I agree with most of Robinson's negative points. I think that all the notions involved are far more difficult than most philosophers have supposed to be the case. I agree with him that Kant is responsible for a great number of muddles in this context. In the first place Kant gives two quite different criteria for analytic truths. The first which asserts that an analytic truth is one in which a judgment is made about something, such that the concept of the predicate is contained, though covertly, in the concept of the subject seems intended to be something of a formal criterion. That is to say that it is a criterion in terms of the sort of thing that one is doing when one makes the judgment in question. When one makes an analytic judgment about something one is in fact analysing the concept expressed by the subject, even though the judgment purports to be about the thing which falls under that concept. The second criterion for an analytic judgment, which asserts with Leibniz that an analytic truth is such that its denial involves a contradiction, seems not so much to give a definition of an analytic truth (as the first criterion might be interpreted as doing) as to give the grounds for asserting that the judgment in question is an analytic truth. That is to say that the second criterion seems to be an epistemological one, not a formal one. And since the distinction between the *a priori* and the *a posteriori* seems to be an epistemological distinction also (in that an *a priori* proposition is one whose truth, while known, could not be known from experience alone) the second criterion for analyticity could

[1] MIND, vol. *LXVII*, No. 267, July 1958.

be set down as an explicit formulation of how we find out that one particular kind of *a priori* proposition is so. (For this reason, Robinson's choice of this notion as a *definition* of necessary truth is wrong. That the denial of a proposition involves a contradiction may be grounds for saying that it is necessary, but it can hardly be what it means to say that it is necessary.)

The Kantian distinction between the *a priori* and the *a posteriori* seems intended, therefore, to be an epistemological distinction, while that between the analytic and the synthetic appears not to be intended to be this, despite one of the criteria given for analyticity. Before going on to consider what Kant says or fails to say about *necessity*, I wish to consider further the first criterion for analyticity.

I have said that it seems intended to be a formal criterion. It must be remembered that Kant applied the adjectives ' analytic ' and ' synthetic ' to judgments, not propositions. Hence if one wanted to give a formal characterization of a particular kind of judgment, one would have to do so by saying what one is doing in making that judgment. For these reasons Kant says that when I make an analytic judgment about X, I make the judgment in such a way that the concept which I apply to X is already included, though covertly, within the concept which I have of X. The judgment " Bodies are extended ", if analytic, is a judgment about bodies such that the concept of a body already includes the concept of extension. It is worth noting that analytic judgments are not judgments about concepts. There has come to be recognized a class of statements about concepts—namely conceptual statements or what Wittgenstein calls *grammatical statements*—but these are not the same as analytic statements. On Kant's view, analytic judgments are definitely about *things*, not concepts, although those things happen to fall under certain concepts. It is by reference to those concepts that an analytic judgment is defined.

If we are to extend the notion of analyticity to statements and still be faithful to Kant's intentions, we shall have to say that analytic statements are statements about something such that . . . . But such that what? The *statement* " Bodies are extended ", if analytic, would have to be a statement about bodies which gives us no information about bodies, because the meaning which we, who make the statement, give to the expression ' body ' already includes the meaning of the expression ' extended '. It is difficult to understand what such a statement would be, and for this and other reasons which I have given elsewhere [1] it seems

---

[1] " Analytic Truths ", MIND vol. lxv, N.S. No. 259, July 1956.

feasible to assert that there are no such things as analytic statements. It should perhaps be pointed out that my reasons for this conclusion are utterly different from those which Quine puts forward in support of his views on analyticity.

To turn to *necessity*. . . . As Robinson points out (p. 292), Kant does not explain what he means by ' necessary truth ', although he connects the notion of necessity with those of universality and the *a priori*. Universality is of course *not* the same as necessity, and I think that it is also evident that to say that a proposition is necessary cannot be the same as saying that it is *a priori*, or for that matter that it is analytic. To say the last is as I have indicated, to give some formal characterization of it, while to say that a proposition is *a priori* is to give an epistemological characterization of it. While each of these may be grounds for saying that the given proposition is necessary, to say that a proposition is *a priori* or analytic cannot *itself* be to say that it is necessary.

In his treatment, Robinson refuses to consider the synthetic *a priori*. Such a refusal may be a little premature. It is at any rate worthy of note that while Kant may be obscure or muddled about what it means to say that an analytic *a priori* proposition is necessary, he is perhaps more clear about the synthetic *a priori*. For propositions or judgments of that kind are said to be such that their truth is a necessary presupposition of experience. Their truth is necessary if experience is to be possible. It may be said, therefore, that for Kant such propositions as " Every event has a cause " are not in themselves necessary, or not necessary *simpliciter*, but only necessary *to* something else. Now, a large part of what I wish to say could be expressed by the statement that in the context of necessary truth the primary sense of the word ' necessary ' is that which it has in the phrase ' necessary to ' (or ' necessary for '). That is to say that the primary use of the word ' necessary ' is its use in that prepositional construction. Moreover, analytic propositions, if any, would be necessary in the same sense. But while I say that that prepositional construction gives the primary sense of ' necessary ', I am very doubtful whether there is any sense to be given to ' necessary ' *simpliciter* in this context. I propose, therefore, to tackle the subject from that angle first.

What, then, is it to say that some truth is a necessary truth? It is, of course, to say that it must be true. But there are a large number of propositions which we should be willing to say, in a given context, must be true, which we should not be willing to count as necessary truths in the philosophical sense. (This fact

may itself be indicative about the concept of necessary truth.) We might, for example, be willing to say, " If Jones says so, it must be true ", where Jones is an authority on the matter in question. But this does not make what Jones says a necessary truth. To say, " If Jones says so, it must be true ", is to say, " It cannot but be true, and that is why—he says so ". This brings out the point that, generally speaking, the use of the word ' must ' has a background of reasons and presupposes that there is such. To say in any context, " It must be true ", is to invite the question ' Why? ' The existence of a background of reasons is brought out explicitly in a formal argument which we conclude by saying, " And so, necessarily . . . . ". It follows from all this that the class of those truths which are necessary in a philosophical sense is narrower than that of those which *must* be true, speaking generally.

The class of truths which are necessary is narrower also than the class of those which are certainly so, even in ordinary usage. To say that something is certainly so is not *eo ipso* to say that it is necessarily so. To say that something is certainly so is to say that one is entitled to be certain that it is so. But from the fact that one is entitled to be certain that something is so, it does not, *eo ipso*, follow that it necessarily is so. If it is certain that p, then there ought to be reasons which give one the entitlement to be certain about the truth of p (although it should be pointed out that saying that there ought to be such reasons does not mean that there is any logical necessity that there should be). But it does not follow from this that these reasons are the sort of reasons which necessitate the truth of p. The reasons which give one the entitlement to be certain that an unsupported and unimpeded body will fall downwards are not such as to necessitate the truth of that proposition. This perhaps is obvious.

It might be argued on the other hand that if something is necessarily so that it must be certainly so. To this it may be admitted that there are undoubtedly contexts in which it would be equally appropriate to use either the adverb ' necessarily ' or the adverb ' certainly '. But what are the reasons for this? Surely they are that one of the things which give one an entitlement to be certain about the truth of some propositions is that the reasons for concluding that they are true are such as to necessitate their truth. To reply that one can be certain of the conclusion only if one is certain of the premises is to ignore the fact that an argument would not be an *argument* unless its premises were taken as certain. To *assert* a proposition in an argument

is to imply that one is not doubtful as to its truth. Never-
theless, although it may be the case that on the occasion on
which it is appropriate to use the adverb ' necessarily ' it is also
appropriate to use the adverb ' certainly ', the converse still does
not follow. Hence it is not right to use the two adverbs as
equivalents. In fact they have quite different functions—
' necessarily ' has a logical function, to stress what follows in an
argument ; ' certainly ' has an epistemological function, to stress
what one is entitled to believe.

In the ordinary way, therefore, to say that some proposition is
necessarily true is to say that it is backed up by certain kinds of
reasons, given which it must be true, and given which its truth is
certain. But a curious feature of necessary truths *as treated in
philosophy*, on the other hand, is that while they may sometimes
(*e.g.* in axiomatic theories) be derived from other propositions, we
should not call these other propositions *the reasons why* they must
be true. While in the case of ordinary arguments leading to a
necessary conclusion we might say that the conclusion *must* be
true if, but only if, the premises are so, or that the truth of the
premises is both a sufficient and a necessary condition of the truth
of the conclusion, it would seem that to say that a proposition is
necessarily true *in the philosophical sense* is to imply that nothing
of a material sort is a condition of its truth. It is, if you like,
unconditional. (Indeed the *O.E.D.* gives ' unconditional ' as a
synonym of ' necessary '.)

Leibniz said that necessary truths are true in all possible
worlds, and this is a way of bringing out the fact that to a neces-
sary truth no particular state of affairs is a necessary condition.
But I think that it should be stressed that to analytic truths, as
traditionally conceived, no *particular* state of affairs is a *sufficient*
condition either. This is brought out by the fact that the sort of
justification of the truth of so-called analytic truths that philoso-
phers have tended to seek is one in terms of the meanings of
words. The truth of so-called analytic propositions does not
follow necessarily from propositions about the meanings of words,
but it does follow *in virtue of* them.[1] Thus, if the proposition
' All bodies are extended ' be viewed as analytic, the fact that its
constituent words mean what they do, given that they are being
used in a language, is a sufficient condition of its truth. But this
sufficient condition does not constitute a particular state of
affairs.

Now the possibility of constructing sentences which look as if
they might be used to state analytic truths is the logical outcome

[1] *Cf.* F. Waismann, " Analytic—Synthetic," *Analysis*, vol. 10, no. 2, 1949.

of the employment of a symbolism.[1] That is to say that if it is possible to construct sentences whose use would be informative, it must be possible also to construct sentences whose use would be uninformative (or, on the other hand, contradictory). Hence, the general possibility of constructing such sentences is a necessary condition of the existence of a language or symbolism. And the possibility of constructing particular cases of such sentences is a necessary condition of the existence of a particular language. Moreover, if such sentences had a use, the truth of the resulting statements would be a necessary condition of the employment of the language from which the sentences are drawn. In consequence I suggest that the sense in which supposed analytic statements are necessarily true is not just that they must be true (for I have pointed out that the philosophical concept of necessary truth is narrower than this), but that their truth is necessary to the employment of the language from which they are constructed, to the employment in a language of the words from which they are made up.

In general, a necessary truth is not just a proposition which is necessarily true (for if this were so, I have argued, there ought to be reasons of a material sort from which it may be concluded that it is necessarily true.) It is, rather, a proposition whose truth is necessary to or for something else. When I turn to the synthetic *a priori* the situation may become clearer. But I am conscious that in saying what I have said I have been making recommendations, or at any rate that I have been attempting to sharpen traditional philosophical conceptions of necessary truth. My excuse is twofold—(*a*) that this conception needs sharpening, since, as Robinson points out, it has never been clear, and (*b*) that while the Kantian tradition implies a dual conception of necessary truth, first, in its application to the analytic *a priori* and, second, in its application to the synthetic *a priori*, my account of the matter brings them both into line.

As is well known, Kant holds that the proposition ' Every event has a cause ' is synthetic *a priori* and therefore necessary ; but he does not hold that its denial involves a contradiction or that the concept of the predicate is contained within the concept of the subject. Again, and perhaps for present purposes more important, it is certainly not true in all possible worlds. For the important feature of the Kantian synthetic *a priori*, as applied to the categories, is that it is concerned with the necessary presuppositions of experience only. Hence, if it is right to say that being necessary involves being unconditional, the Kantian

[1] *Cf.* Wittgenstein, *Tractatus* 4. 4611.

synthetic *a priori* judgments are in the paradoxical position of being conditionally unconditional. For, they are true, subject to no conditions except the condition that they apply only to experience. Hence, that *there is* experience is itself a necessary condition of their being true *of* anything. But in what sense are they unconditional at all? In what sense are they necessary? It cannot be maintained that their truth, if they are true, is necessary to the employment of their constituent expressions, as is the case with so-called analytic truths. Yet, nevertheless, there is a certain similarity between them and the latter, in that their truth is certainly necessary to something else.

To continue with the same example—Kant's argument in the 2nd Analogy is most plausibly represented as designed to establish a *conceptual connection*. His contrast between our experiences of the ship and the house, when the series of experiences of the one is objectively determined while the series of experiences of the other is subjective in order, is designed not so much to show that objective events always involve regularity (for how could a hypothetical example show this?) as that our concept of an objective event involves the concept of regularity. For Kant, this means that the concept of *cause* is connected with or involved in the concept of an experienced event, or alternatively, that the connection between the concept of an event and the concept of cause is necessary to the concept of experience which we have. But this connection is not such that the one concept includes the other ; we cannot *define* an ' event ' by reference to the concept of a cause, nor did Kant wish to do so.

It is indeed important to see that conceptual connections are not all of the same kind. For example, it could be argued that the concept of general acceptance is connected with that of accept-ability, in the sense that if a thing is worthy of acceptance there is a presumption that people will, in general, accept it. But it would be a mistake, now well-known, to make the connection a necessary one. In the case of the concepts of cause and event Kant argues with a certain plausibility that the connection between them is a necessary one *relative to the concept of* experience. One of the facts of fundamental importance about Kant is his recognition of arguments of this sort, those, namely, which he called ' transcendental '.

Given what I have said so far, it follows that the statement " Every event has a cause " may be considered necessary in the sense that the truth of " X has a cause " is necessary to the truth of " X is an experienced event ". Of course, on the face of it the statement *sounds* as if it is meant to be contingent, and a sense

could no doubt be given to it in which it is so. Consequently, if it is to be considered necessary it must be given a rather special use, and that use is to assert a conceptual connection. But that someone is using the statement in that way and that the connection which is being maintained is what it is can be appreciated only by considering the arguments which lead up to it. It can be appreciated only by considering Kant's actual arguments that he intends the statement " Every event has a cause " to say that the connection between the concepts of cause and event is necessary to our employment of the concept of experience—given that he does so intend it.

I have so far distinguished two possible candidates for the title of ' necessary truth '—one, a statement about things, such that, if there is any such candidate, its truth is necessary to the employment of the language from which it is derived, the second, a statement about concepts, such that its truth is necessary to the employment of some other concept or concepts. One thing which distinguishes the two cases is that the first purports to be a first order truth about objects, while the second is a second-order truth about concepts. Another thing which distinguishes them is that the first is such that its truth is necessary to the employment of the expressions used or to the employment of the concepts exemplified in those expressions, while the second is such that its truth, namely the truth about a connection between concepts, is necessary to the employment of *some other* concept. They are united in that each is, in its own way, necessary to a particular mode of thought (*i.e.* to a particular use of concepts). Hence if their truth is necessary to something particular it is to a mode of thought, not to some particular state of affairs. The *differences* between the ways in which these two kinds of statement are necessary may well be more important, philosophically speaking, than the similarities between them, but I would stress again that these differences are not, strictly speaking, differences in their necessity.

It might be asked whether there are any particular statements whose truth is necessary to *any* mode of thought whatever, to the employment of *any* concept whatever, or (if I may use a phrase which I have so far avoided) to *any* language-game whatever. The truth of *some* analytic statements would, if there were any, be necessary to any language-game ; that is to say that the use of *any* language-game presupposes that *some* analytic statements should be true if there are any at all. But it is not, of course, the case that the truth of any *particular examples* of analytic statements is necessary to *any* language-game. As regards conceptual

statements the situation is obscure. The traditional laws of thought come to mind here, for they seem all-pervasive. I wish, therefore, to consider them briefly.

If the laws of thought are not to be construed as rules, but as statements, one way of taking them is to interpret them as saying something about concepts connected with the use of expressions. For example, the Law of Excluded Middle can be taken as saying " All statements are *either* true *or* false ". To say this is at least to say something about the connection of the concept of a statement with the concepts of truth and falsity. It would, I think, be ridiculous to maintain that this connection is a factual one ; to attempt to verify the Law of Excluded Middle by surveying numbers of different statements would manifestly be to misunderstand the point. So there is some justification in saying that *qua* statement it is necessary. But I do not think that it would be right to view it as analytic. It would be wrong, that is, to define ' statement ' in terms of truth-or-falsity.

Wittgenstein has a discussion of a connected point at *Philosophical Investigations*, i. 136. There he takes it that we *might define* a proposition as *whatever can be true or false*, and goes on to ask in what sense we might talk of determining what a proposition is by saying that it is what fits the concept ' true '. But he goes on to say that so to define a proposition is to say no more than that we call something a proposition if we use ' true ' or ' false ' of it. Or, in other words, the use of the expression ' proposition ' and the expressions ' true ' and ' false ' are both moves in the same language-game. But this seems to bring out the point also that it is the nature of the language-game which determines whether it is necessary that *either* the expression ' true ' *or* the expression ' false ' should be applicable to any given proposition. Thus, the connection between the concept of a proposition or statement and that of truth-or-falsity is necessary to the employment of the concept of a particular language-game, and *not* to the employment of those concepts themselves. The truth of the proposition " All statements are *either* true *or* false " is necessary to the employment of the concept of a particular language-game, namely that concerned with the making of assertions. Because of this it would be wrong to say that the Law of Excluded Middle is analytic.

It might be objected that the proposition " Two plus two is four " is analytic, and yet surely its necessity is relative to a particular language-game—namely using numbers in, *e.g.* counting. There is much that might be said here, but I am inclined at any rate to say that if " Two plus two is four " is to be looked on

as a proposition at all, and one whose truth is necessary at that, it is in many ways on a par with the Law of Excluded Middle. It may be regarded as stating a connection between concepts, a connection which is necessary to the employment of the concept of a particular language-game, *i.e.* the use of numbers. In this it differs from " All bodies are extended ", where the truth of the proposition is not necessary in relation to a given *kind* of use if expressions, even if it is necessary to the employment of the constituent expressions. In other words, " Two plus two is four " is, if necessary at all, necessary to the employment of the concept of a particular language-game, *i.e.* to the employment of the concept of a particular *mode* of using expressions ; while " All bodies are extended " is not necessary to the employment of the concept of a particular mode of using expressions, but to the employment of the particular expressions or concepts of which it is constituted. For these reasons I would not on any account treat " Two plus two is four " as analytic, whatever else it is.

I conclude, likewise, that if the Law of Excluded Middle is to be considered as necessary at all [1] it is because it is necessary to the employment of the concept of a certain language-game. But it is worth noting that all that is being claimed in this is that it is necessary to the employment of a *particular* language-game. It might, therefore, be asked whether more generality could be obtained by considering it as necessary to the employment of the concept of a language in general. In the *Philosophical Investigations* Wittgenstein refuses to allow so much, by implication at any rate, on the grounds that there is nothing common to all applications of the expression ' language ' (see *P.I.* 65 ff.). But his arguments are not altogether convincing. As is well known, in the early sections of the *Philosophical Investigations* Wittgenstein is chiefly concerned to show that the meaning of an expression is not its reference. He argues towards this conclusion *via* the consideration that those things which are called a ' game ' or a ' language ' are not very similar. But what is to count as similarity here? Is, for example, similarity of function one of the criteria? In section 11, Wittgenstein maintains that the functions of words may be as diverse as the functions of different tools, and he adds " And in both cases there are similarities ". But when he insists that the things which are called a game may have only a family-resemblance, he seems to be implying that we should consider these different games merely from the point of view of, for example, the moves in which they consist. That is to say

[1] As in the case of " Every event has a cause " there is a sense in which it might be viewed as contingent, and even false.

that he imagines that the word ' game ' might be defined merely
by reference to what people actually do in games, without refer-
ence to, for example, conventions and intentions.

In the context it is, of course, right that he should do this,
for his main object is to combat the view that the meaning of
any expression can be given by reference to an object to which it
refers. And certainly one could not define ' game ' or ' language '
in that way. But it does not immediately follow from this that,
if the things which fall under a concept differ or are not exactly
similar, the concept itself has no principle of unity. If it has,
we should surely be able to say what this is and what is necessary
to its existence. Pointing to examples like that of language-
game two of the *Philosophical Investigations* is not sufficient as a
counter-argument against the view that the truth of the Law of
Excluded Middle is necessary to the concept of a language.
That is to say, it is not enough to point to a language-game in
which, *prima facie*, statements, to which the Law of Excluded
Middle might be applicable, do not occur. For what enables us
to say rightly that this *is* a language-game, and not just an exer-
cise with noises or even signs, like the making of signals? I do
not think that we should count something as being a language
unless something is capable of being *said* in it. If we allow that
the utterance of words like ' Slab ' and ' Beam ' is a case of
saying something I am inclined to think that we are allowing the
concept of a statement to be already involved, at any rate
implicitly. There is this much truth perhaps in the Fregean
claim that only in the context of a proposition does a word have
meaning.

However this may be, it still remains true that *if* the Laws of
Thought are to be viewed as necessary they are necessary to the
employment of a single concept only—namely, that of a language-
game, or, if greater generality be permitted as perhaps it may be,
that of a language. But it follows that they are not necessary
to any concept whatever. And I know of no other promising
candidates. It may perhaps be concluded, therefore, that all
necessary statements are necessary to something particular, to the
employment of a particular concept or a particular mode of
thought, and not to any mode of thought whatever.

My general conclusion from this examination of candidates for
the status of necessary truth is that the philosophical notion of
necessary truth is that of a hypothetical necessity of a particular
kind—in that examples of necessary truths are necessary only
to a particular mode of thought. Opposed to such truths are
those which are not necessary but only possible with regard to any

particular mode of thought. Hence, the concept of necessary truth should, strictly, not be opposed to that of contingent truth. Those statements whose truth is not contingent are those which, as Aristotle said, are true in themselves (καθ' αὑτο), and whether we should say that either analytic or conceptual truths are true in themselves is a matter for doubt. However that may be, there do seem to be grounds for distinguishing between the concepts of statements which are (a) necessarily true, (b) true in themselves, (c) analytic, (d) a priori, and (e) non-empirical. Those statements which fall under these concepts may sometimes (though not necessarily always) overlap, but it remains true that the concepts themselves are not to be identified.

University of London

VII.—PARTIAL ENTAILMENT AND THE CAUSAL RELATION

By R. G. Robinson

THE term "partial entailment" is not a new one, but to my knowledge there is nowhere in the literature an explanation of exactly what is meant by it. Perhaps nothing precise has indeed been meant or perhaps the meaning is unavoidably vague. Whatever its past meanings have been, I am using it to stand for a logical relation which, so far as I know, has not hitherto been recognized, and one that is of the utmost importance in the analysis of causation. Indeed, without it no such analysis is likely to be adequate ; for it explains that aspect of a causal connection that is best described as one of necessity, yet which is not an instance of logical necessity.

For everyone who holds that being black is not a necessary property of ravens the proposition that this raven is not black is not contradictory, though it may be false. Moreover a similar argument can be used with reference to the other properties that ravens are commonly found to possess, such as being 18 inches tall, having a bill $1\frac{1}{2}$ inches long, and having a liking for grain. [Although these properties may not be the best criteria for ravens, for the sake of the argument in this article they are assumed to be.] "This raven is 5 feet tall ", though probably false, is also not contradictory. Similarly the other two properties can be absent, yet the bird be a raven. What cannot be said without contradiction is that this raven is not black, is not 18 inches tall, has a bill two feet long, and does not like grain. The generalized argument runs : if a, b, c, d, be the usual properties looked for in a thing that is called an x, then " x is not a ", " x is not b", " x is not c ", and " x is not d " are synthetic statements, whereas " x is not a and not b and not c and not d " is logically false. The relation that underlies this peculiarity is the one I am calling partial entailment.

But the relation is a little more peculiar than appears in the generalized expression, for a synthetic proposition that denies a single property of a thing may not be *simply* synthetic. Let us suppose there are only four properties that are usually found in a thing of a certain kind, and that the absence of one of the properties, or even of two of them, in any particular case will not disqualify the thing from being of that kind. There are thus nine combinations of properties which will make the thing that kind of thing. Let them be expressed as a disjunction of conjunctions:

abc v abd v bcd v ab v ac v ad v bc v bd v cd.   Thus when any one of these combinations occurs the thing will be of the kind whose criteria are given in this expression.   If a thing is x, then it possesses at least two of the four properties.

The foregoing is clear and straightforward, with no suggestion of a strange relation or the need for one.   But how is a proposition of the form " x is not a ", in which only one property is mentioned, to be classified ?   Since a is not by itself an essential property, such a proposition is not analytically false.   However, if a is denied at the same time as the other three properties are denied, then the conjunction of the four denials is logically false.   Hence, a is not entailed, for if it were, then its denial would be contradictory ;   and this is not the case.   Yet neither is it not entailed since it could then be denied under any circumstances without contradiction ; again this is not the case.   Thus in spite of the fact that the definition of x can be given in terms of conjunction and disjunction, with regard to any single property a new element is involved that is not a simple matter of conjunction and disjunction.

Referring to the instance of ravens, we have " ravens are not black " as a synthetic proposition, but with the underlying feeling that there is another factor involved.   Analysis reveals that the extra factor is the feeling that although the proposition is synthetic and so possibly true, it nevertheless is the case that it could not be true, logically could not be true, under certain specifiable conditions.   For example, when conjoined to " ravens are not 18 inches tall and ravens do not like grain," it cannot logically be true ;   nor can it be logically true when conjoined to the conjunctions of the denials of the other possible combinations of properties.   Because of this feature, single properties are neither entailed nor not entailed, neither necessary nor non-necessary, neither essential nor non-essential to the things of which they are properties.   Consequently they may reasonably be described as partially entailed properties, and the relation obtaining between them and the things to which they belong as partial entailment.

However a qualification must be made, for not all single properties are partially entailed.   That is to say, some properties are genuinely contingent.   For instance a bird having a single eye would not necessarily be disqualified as a raven, and roses need not be any special colour.   The question arises then, " How are contingent properties to be distinguished from those that are partially entailed ? "

The way to distinguish one kind of property from another is to think of how you yourself regard the property.   If you think

that in the absence of others it could also be absent without affecting your classification of the thing in question, then it is a contingent property. But if it is one of the set of properties, at least some of which you consider necessary to a thing's being that kind of thing, then it will be a partially entailed property, if it is not analytic. For example, a person who thinks nothing strange in the idea of non-black ravens might mean either of two things. In the first place it might only signify that he does not think of blackness as anything but a contingent property of ravens. But in the second place, he might mean that being black is one of the partially entailed properties of ravens ; and this means that if too many of the other partially entailed properties disappear then the bird is not to be considered a raven.

That is to say, if under 'any combination of properties the presence or absence of another property a makes no difference to your classification of the thing in question, then a is regarded as a contingent property. But if too many other specifiable properties are missing and the presence or absence of a makes the difference between classifying the thing in one way instead of another, then it is a partially entailed property, if it is not an analytic one.

Digression : It is this relation that causes disputes about the contingency or necessity of propositions. For the fact is that we do not all of us hold as partially entailed the same set of properties for any particular thing. It is thus possible for one person to assert that a proposition is undoubtedly necessarily true, while another asserts that it is undoubtedly contingent, and the fact of the matter be that owing to the property referred to being really partially entailed the proposition about it is neither. Yet even this is not quite right. For the proposition might indeed be analytic or it might be contingent. Whether it is one or the other will depend not on any empirical facts but on how a particular person uses the language in such cases, i.e. how he thinks about them, what his conceptual scheme is like. And if the property is really partially entailed, the fact that it is not generally recognized to be so—and in my opinion never has been recognized until now—causes the ambivalence shown towards it, an ambivalence that is manifested by calling the proposition analytic, or synthetic, or, by some people who have never been able to explain exactly what they meant by the term, synthetic a priori.

However, anything that has so far been said will not explain the necessity of the causal connection, since the argument has been confined to the co-existence of properties in a particular

kind of thing.   Although these properties are often called causal properties, not many people would seriously argue that any one of them provides us with a genuine instance of causation.   No one maintains that the blackness of a raven causes its being 18 inches tall, or the converse either.   It is not argued that the size of a tree causes the shade of greenness it possesses, although if such a relation were established we might well describe the one as the cause of the other ; though if this were the case, I fancy the size would be taken as the cause and the colour the effect, rather than the colour the cause and the size the effect.   Whether such properties as we have been considering are inaptly called causal or not—I think they often are—causal connections are usually, and in ordinary affairs always, attributed to a succession of events, a sequence or series of events.   In the sequence the earlier is always called the cause of the later event.

I suppose a word should be said about events since it is extremely difficult if not impossible to give an adequate definition of " event ".   For the purpose of showing how the relation of partial entailment is relevant to the causal relation it is not, fortunately, necessary to have a precise definition.   One situation that is usually described as causal is that in which when a switch is turned a light comes on.   Now the turning of the switch can be reasonably described as an event, and so can the lighting of the lamp.   Setting fire to a paper and its reduction to ashes can be considered as two events.   Putting fertilizer on a plant and subsequent features of its growth may be two events.   In any one of these cases both events can be considered as different aspects of a single somewhat complex event, but there is no advantage in doing so if one wants to clarify the causal relation. Each can be distinguished from other states of affairs that, however much they may be plausibly thought of as composed of numerous smaller events, are not considered events themselves. For instance, although in one sense a dog is composed of successive states of flesh and bone, in another sense the dog is not an event at all.   We do not say that the dog as it was yesterday caused the dog as it is today, and one of the reasons is that neither seems to be an event, although we might well say that the food the dog ate yesterday caused its discomfort today.   So without attempting a metaphysical inquiry into events for the purpose of reaching a general statement of criteria of them, we can nonetheless have a fairly clear idea of what events are in the ordinary world ; and no more is needed for the understanding of how the relation of partial entailment can account for the aspect of necessity which is so universally found in causation.

Being a raven partially entails being black, being 18 inches tall, having a beak $1\frac{1}{2}$ inches long, and liking grain. Such a notion may be described in terms of different levels of language, or of concepts, or of ideas, or of different complexity in ideas, some being simple and at least one of them complex. But this matter can be left to those people who want to get their technical jargon accurate and complete, or those who are interested in the metaphysical nature of properties and the things of which they are properties. The co-existence of certain properties leads us to refer to the set by a short term which, acquiring meaning of its own, will have repercussions in the logic of our thought or language. This of course is not a psychological account of how we come to know about ravens and their distinctive character- istics. The logic of the matter, however, is brought out by a consideration of what being a raven entails, or, as I think, par- tially entails ; and this suggests that we find out what properties co-exist and, in order to refer in a short way to a number of co- existent properties having the same spatial and temporal co- ordinates, use a term that comes to have connotation in terms of these very properties.

But being a light does not so clearly depend on the partially entailed property of the turning of a switch, yet such a turning would ordinarily be called the cause of the light's coming on. How does the notion of partial entailment apply to sequences of events as opposed to the co-existence of properties ?

Consider the case of turning on a lamp. The switch is turned and the lamp lights, it is turned again and the lamp goes off. Let us confine the discussion to the first turning, the one that would normally be called the cause of the lamp's lighting. Although the switch only completes the electric circuit and although certain mechanical and electrical conditions must obtain, these considera- tions are normally ignored, so that the causal relation is asserted simply between turning a switch and the lighting of the lamp. There are two concepts that have to be examined, one of them being " switch " and the other " lamp ". It is true that if we ask for a switch in an electric shop we will be handed a device that is connected to nothing and, if left unconnected, might serve as a child's toy. What makes it a switch is the fact that it can be included in a mechanical and electric arrangement of wires and power supply such that in one position certain things happen and in another certain others do. If there were not at least two possibilities, each dependent on different positions of the switch, it would not be a switch. If the points on a railway line did not have the effect of leading a train to one track or

another, there would be no point in calling the points a switch, as they are in some countries. For those people who prefer to have the same idea presented in terms of concepts, the concept " switch " is a complex one involving the concepts of " change " and " effect " and a system or unity, such that if a change in the system has no effect then the switch is not really a switch. In other words, if we turn what would ordinarily be called a switch, and if nothing happens even though wiring, power supply, and lamp are all in working order, then we have not really turned a *switch* ; it is a contraption whose position has been altered, but it is not a switch. Similarly, an oddly shaped piece of glass, with metal at one end of it and a thin tungsten wire in the vacuum enclosed, is not really a lamp if it will not light under any conditions. If such an arrangement of metal and glass or perhaps of some other material, were to light while it was lying on the table unconnected to anything, it would still have lighted. We can tell the difference between a lamp that is lit and one that is not ; it so happens that the ways of getting them to light are, to our present knowledge, severely limited. But if the metal and glass were never to light up, regardless of how they are arranged in themselves or how connected to anything else, then there would be no lamp.   For the concept of lamp includes the concept of lighting.

In our particular instance there is a further consideration. Switches are not all of them light switches.   There are rail switches, switches for turning on electric irons, for operating lifts, and a great number of other kinds.   In each case the reason any particular modifier of the term is used is that the turning of switches can be related spatially and temporally with certain other events.   If whenever the switch on my desk-lamp were turned the overhead light came on, I would not call that switch the desk-lamp switch.   If, in order to put the desk-lamp on, the wall switch had to be turned, then that one would be called the desk-light switch.   If there are two switches on the panel in a lift and one operates the light and the other the lift—it is a strange lift—then one will be called the light switch and the other the lift switch.   It should now be clear that the turning of a switch is associated by us with certain other changes in the environment.   But because of this possibility of association the turning of a light switch, not merely of a switch, partially entails the coming on of a light.   The turning of an electric fire switch partially entails the coming on of an electric fire.   The coming on of a light partially entails the turning of a switch, just as the lighting of an electric fire partially entails the turning on of an

34

electric fire switch. "Partially entails" only, because the relation is not a necessary one, and also because from a single failure in a relationship the conclusion will not logically follow that there was not a switch, or not a light, or not a fire, as the case may be. That is to say, concerning light switches for instance, one of the partially entailed properties of light switches is that they put on lights when turned ; but it is not a necessary property, since failure to do so will not mean that they are not light switches, though constant failure will. If the fire fails to get hot when the switch is turned, there might still be a fire and there might still be a fire switch, for the relation between them is only partially entailed, although if this piece of metal never gets hot it is not a fire and if the switch never turns it on then it is not a fire switch.

Now, in my opinion, the necessity of a causal relation is to be attributed to the relation of partial entailment that obtains between the things said to be causally connected (or between the concepts that stand for them, or the words that can be used to refer to them : the description is a matter of metaphysical preference, perhaps based on aesthetic delight). For causal necessity is not full analyticity ; but it is necessity, and not merely contingency. However we do not look for the necessity in a peculiar relation in the world, but rather in the way we think about the world. Because in thinking of certain things as causally related we do relate them by means of partial entailment, by means of their partially entailed properties, we find that the necessity associated with causation can be accounted for in terms of this relation. Ravens must be black, but they do not absolutely have to be. Light switches necessarily turn on lights, though they do not absolutely have to. Expressing the distinction in the old terminology of necessity and contingency, of analytic and synthetic propositions, is difficult because we have the meanings of these words so fixed as to exclude the kind of necessity involved : causal necessity. But some concepts are in fact related in the way that has been shown, the way I am calling that of partial entailment, in which there is a kind of necessity involved which is not analytic necessity. It seems to me to explain what has been for the past 200 years the most puzzling feature of causation.

An example of its effectiveness can be found in the way it enables a distinction to be made between a universal of fact and a universal of law. A non-necessary statement of the form "All A's are B" can be given two interpretations : "If anything is A it must also be B", and "If anything is A it will also as a matter of fact be B". The problem lies in giving meaning to "must", to show that given the constant conjunction of A and B it is

nevertheless possible to distinguish the case of A and B that are causally related from the case of A and B that are not. The solution, I think, is that the statement expresses a universal of law if the components are partially entailed, and it expresses a universal of fact if they are not partially entailed. If we find reason to think of A as partially entailing B and *vice versa*, we will no doubt think that " All A's are B " asserts a law ; whereas if we see no reason to relate them in this way we will regard the statement as an expression of fact only. Though the details are boring, they can be worked out by any interested person who understands the relation.

The University of Witwatersrand

THE RE-DISCOVERY OF THE TOPICS

Professor Toulmin's Inference-Warrants

THE view is gaining some currency that laws of nature may be profitably considered as inference-licences. Professor Ryle advanced it in the metaphor of inference-tickets (R. 121).[1] Professor Toulmin, on his own admission, applied Ryle's ideas to the elucidation of physical science in his *Philosophy of Science* (T. 260). Much the same idea was put forward by Schlick, who told Professor Popper that he owed it to Wittgenstein (P. 37, n. 7). Ramsey also said much the same thing (RS. 241). More recently, Toulmin has launched what amounts to a full-scale exposition of inference-warrants in his study on the *Uses of Argument*.

This development, particularly in the form it takes in this last book, has many similarities with the analysis of the Topics in medieval logic. The resemblance is so close, as I hope to show, that it appears we are witnessing something of a re-discovery of the Topics. To make manifest the resemblance I shall describe first what Toulmin has to say about inference-warrants, particularly in the third essay of his book. I shall then present the medieval doctrine on the Topics as found in the tradition that stretches from Abelard to Albert of Saxony in the fourteenth century.

Professor Toulmin's Warrants

Concerned with what he calls " working logic " rather than with the " idealized logic " of the formal logicians, Toulmin prefers to deal with examples expressed in concrete terms. Thus for the analysis of the inference-warrant he gives the following examples :
1. Harry's hair is red, so it is not black.
2. Petersen is a Swede, so he is almost certainly not a Roman Catholic.
3. Harry was born in Bermuda, so he is a British subject.
In each of these we find what is called a *Datum* and a *Claim*. In the second part of each statement we have the " *claim* or conclusion whose merits we are seeking to establish (C) ", and in the first part " the facts we appeal to as a foundation for the claim . . . our *data* (D) " (T. 97).

However, of each of these statements we may still ask, how we get from the Datum to the Claim. To answer this " we must bring forward not further data, for about these the same query may immediately be raised again, but . . . rules, principles, inference-licences, . . . general, hypothetical statements which can act as bridges, and authorize the sort of step to which our particular

---

[1] These references are to the books listed at the end of this paper.

argument commits us " (T. 98). These are called *Warrants* (W). The following statements are offered as warrants for the three arguments above :

1*. If anything is red, it will not also be black.

2*. A Swede can be taken almost certainly not to be a Roman Catholic.

3*. A man born in Bermuda will be a British subject.

In general then we get the following diagram for the " layout of arguments " (T. 99) :

$$D \longrightarrow So\ C$$
$$|$$
$$Since$$
$$W$$
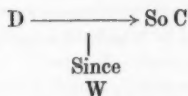
However, for a full picture of the layout we must amplify the diagram. We may be asked " why *in general* this warrant should be accepted as having authority ", and to this we will produce " other assurances, without which the warrants themselves would possess neither authority nor currency ". Such assurances are called the *Backing* (B) of the warrants (T. 103). As examples of such, he cites for 2 and 3 :

2**. The proportion of Roman Catholic Swedes is less than 2 %.

3**. The statutes and other legal provisions governing the nationality of persons born in the British colonies.

No backing is given for the first argument, presumably on the ground that the warrant itself is " self-authenticating " (T. 98). The Backing, it is noted, can be expressed in the form of categorical statements of fact, like D, and unlike W (T. 105-106).

Furthermore, we may want to qualify the degree of force which the data confer on the claim in virtue of the warrant, which is accomplished by such " modal qualifiers " as ' necessarily ', ' probably ', ' presumably '—which will be indicated by Q. It may also be necessary to state the conditions for the exception or rebuttal (R) of the warrant as indicating the circumstances in which the authority of the warrant would have to be set aside, *i.e.* statements about its applicability (T. 100-102).

With these amplifications our diagram will now appear thus (T. 104) :

$$D \longrightarrow So,\ Q,\ C$$
$$|\qquad\qquad |$$
$$Since\qquad Unless$$
$$W\qquad\qquad R$$
$$|$$
$$On\ account\ of$$
$$B$$
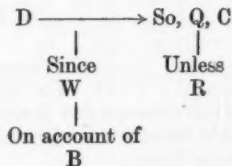
For concrete illustration Toulmin returns to the case of Harry who was born in Bermuda (T. 105). However, it should also hold for the example from geometric optics which he exploited in his *Philosophy of Science* (T* 24-25). Applying these distinctions, we would then get something like the following :

    4. (D) The sun is at an elevation of 30° and a wall is 6′ high,

       (C) So the wall will cast a shadow of 10′ 6″.

4*. (W) By the Principle of the Rectilinear Propagation of Light, the sun's rays can be represented as a straight line and the shadow of the wall can then be computed by trigonometry.

4**. (B) Particular cases in which this method of computation has been shown to work—Newton's " rendering a proposition general by induction ", which Professor Toulmin interprets as " using our observations of regularities and correlations as the backing for a novel warrant " (T. 121).

    4. (Q) In this case, a statement of the margin of error that would be tolerated.

    4. (R) The exceptions, *e.g.* diffraction, refraction, in which the warrant (4*) does not hold.

This is a more elaborate analysis than Toulmin made in his earlier work. Yet as Professor Cooley has indicated, it would still seem to be an over-simplification when applied to the analysis of even so simple a law as that of Snell ; since it is not immediately clear whether such relevant items as the success of comparable laws, its tie-up with the wave-theory of light, and its prediction record are to be counted as D, W, or B (CL, 312-315). However, this is not immediately relevant to our concern of the similarity of Toulmin's analysis with the Topics, to which we can now turn.

The Topics

The examples make it clear that Toulmin is primarily concerned with arguments which derive at least some of their argumentative force from relations of meaning among the non-logical words—*e.g.* Swedes and Catholics, native Bermudans and British citizens. In fact, he declares that the backing of an argument is " field-dependent" in that it " varies from one field of argument to another " (T. 103). This is to say, in terms of the medieval logical analysis, that he is concerned with material rather than with formal consequence. ' *Formal* ' in this connection has to do with the syncategorematic terms, such as the connectives, ' and ', ' or ', ' if . . . then ', ' not ', and the quantifiers ' all ' and ' some ', whereas ' *material* ' refers to the categorematic terms (AS. f. 24rb). The logical study of material consequence, *i.e.* of logical consequence that depends in some way upon the categorematic terms, was for medieval formal logic primarily the study of the Topics.

The medieval Topics tradition, which stems from Boethius, drew a two-fold distinction in a Topic:

(1) The Topical Maxim, or rule according to which one can make an inference—the *maxima propositio*.

(2) The Topical Difference, or the ' matter ' from which the Maxim is composed—the *differentia maximae propositionis*.

What these two are and how they function may be seen from the analysis of a typical Topical argument:

5. If it is a man, it is an animal.

(Abelard regularly cites Topical arguments as conditional propositions with ' if ', although later, *e.g.* in Peter of Spain, Albert of Saxony, it was customary to cite them as two propositions connected by ' therefore ').

> 5*. *Topical Maxim* : Of whatever the species is predicated, so is the genus (*De quocumque predicatur species, et genus*).
>
> 5**. *Topical Difference* : ' Man ' is related to ' animal ' as species to genus.

The second of these is called the ' Difference ' because it differentiates one Topic from another.   So in this case we have the Topic of Genus and not the Topic of Definition, Whole, Part, Opposites, etc.   Both Maxim and Difference are called ' Topics ', Peter of Spain writes, " because both confer validity (*firmitatem*) on an argument " (PS. 5.07).   A Topic itself is most briefly described, after Cicero, as a ' basis of argument ' (*sedes argumenti*).

Now I would suggest that in the Topical Maxim and Difference we have the traditional logical counterparts of Toulmin's Warrants and Backing.   The Maxim performs the same function as a warrant and to describe it as " a confirmatory rule that proves a consequence ", as Albert of Saxony does (AS. f. 33ra) is nothing Toulmin should object to, except perhaps for its departure from the sound of ordinary language.   The Difference by its name indicates that diverse Maxims rest on different relations.   Toulmin himself cites the genus-species relation as backing for a warrant, when he says that ' All whales are mammals ' as a warrant for an argument between two singular statements has as its backing, ' The class of B's includes taxonomically the entire class of A's (T. 112).

Furthermore, the first of Toulmin's examples cited above (*i.e.* 1) is itself an instance of one of the traditional Topics, that namely from opposites.   The medieval logician, however, would push the analysis of it further along the same lines than Toulmin seems willing to go, since as we have seen, he apparently holds that its warrant, in being " self-authenticating ", needs no backing.   A Topical analysis of this example would appear something as follows :

> 1'. If Harry's hair is red, it is not black.
>
> 1*. *Maxim* : Positing one contrary removes the other from the same.

1**. *Difference* : ' Red ' and ' black ' are related as contrary opposites. This involves recognizing that they are opposites, sorting out what kind of opposites they are (contrary, relative, privative, contradictory), so as to know that they are such that both cannot be present at once in the same subject ; hence that the Maxim applies.

The Formal Interest of the Topics

Enough has been said to show the similarity between Toulmin's analysis of warrants and backing and medieval Topical logic. However, a short word may be in order about the difference between them, even though this is to broach what is perhaps the main thesis of Toulmin's book, *i.e.* the over-formalization of logic ; and this is indeed another and bigger story.

The Topics during the great period of medieval logic, say from Abelard to Paul of Venice, were studied primarily for their formal interest. In fact, a Topic as a ' material consequence ' is to be understood with reference to a ' formal consequence '. Thus the Pseudo-Scot writes that whereas a formal consequence is " one that holds for all terms when there is a similar disposition and form of the terms," a material consequence is " one which does not hold for all terms retaining a similar disposition and form but with a variation of the terms " (S. 287B). However, and this provides the centre of formal interest in the Topics, a material consequence " can be reduced to a formal consequence through the assumption of a necessary proposition " (S. 288A). How this is done can be seen from our example No. 5 :

5. If it is a man, it is an animal.
If we add to this the proposition :
5'. Man is a species of animal
we get the following ' formal consequence ' :
5″. If man is a species of animal, then if it is man, it is animal.

This consequence is ' formal ' because it holds good in virtue of the form regardless of the terms that realize it. It can accordingly be formalized with variables. Interpreting the relation in terms of class-inclusion and class-membership, 5″ then becomes the following thesis of the logic of classes :

5″*. $\alpha \subset \beta : \supset : (x) \cdot x\epsilon\alpha \cdot \supset \cdot x\epsilon\beta$

In effect, we now have an abstract statement of the Topical Maxim 5*. Taking this and the Difference (5** or its equivalent 5') as asserted, we may detach and assert 5 by itself. In this way the whole Topical structure can be viewed as an inference-scheme, as in fact Abelard does.

Abelard throughout his long study of the traditional list of Topics (A) is mainly concerned to discover those which exhibit such formal

relations and to distinguish the corresponding necessary Maxims from those which are probable only. John of Salisbury, while admiring his logical acumen, complains that he is interested only in necessity (J. III. 6). Toulmin sounds an echo of this complaint, though with something less than admiration for the formal possibilities. But that more has been done in the analysis of warrants and backings than Toulmin seems to realize is clear from even so brief a view of the medieval Topical tradition.

References

A Abelard, *Dialectica*, ed. L. M. De Rijk, Assen, 1956.
AS Albert of Saxony, *Perutilis Logica*, Venice, 1523 (courtesy of the Bodleian).
CL J. C. Cooley, " On Mr. Toulmin's Revolution in Logic " in *Journal of Philosophy*, lvi (1959), No. 7, 297-319.
J John of Salisbury, *Metalogicon*, ed. C. C. J. Webb, Oxford, 1929.
P K. R. Popper, *The Logic of Discovery*, London, 1959.
PS Peter of Spain, *Summulae Logicales*, ed. I. M. Bochenski, Rome, 1947.
R G. Ryle, *The Concept of Mind*, London, 1949.
RS F. P. Ramsey, *The Foundations of Mathematics*, London, 1931.
S Pseudo-Scot, *Quaestiones super libros Priorum Analyticorum*, in Duns Scotus, *Opera Omnia*, ed. Durand, Lugduni, 1639, T.1.
T S. E. Toulmin, *The Uses of Argument*, Cambridge, 1958.
T* S. E. Toulmin, *The Philosophy of Science*, London, 1953.

OTTO BIRD

University of Notre Dame,
In   .1a

DRAY ON HISTORICAL EXPLANATION

In his critical notice (MIND, April 1959) of William Dray's recent book, *Laws and Explanation in History*, P. F. Strawson gently takes the author to task for stopping short " at crucial points in his own positive account " of historical explanation, for his being " willing rather to leave us with a mystery than to uncover grains of truth in the disputed thesis," *i.e.* the " covering law theory ". With this criticism I fully concur. Strawson's subsequent remarks, however, only partially uncover those " grains of truth " and tend to sustain rather than clear up some of the mystery of historical explanation.

The particular mystery I have in mind is the sort of evidence which guides the historian's judgment when he is trying to ascertain why an event happened. Suppose that we agree with Strawson that such judgments are not typically assertions of necessary connection, *i.e.* that the historian does not typically assert that between the event and the circumstances cited to explain it there was a necessary connection. Instead, we agree that the historian's typical claim is that the event was probable in view of such and such conditions ; that, in Strawson's words, " it is entirely natural that it should have happened as it did " (p. 267). Suppose we further agree with Strawson, as against Dray, that " analogies with other cases " and " truisms about human behaviour " are " at work in guiding the selection of facts and their arrangement into explanations which satisfy the judgment " (p. 267). The mystery remains as to how such analogies and truisms guide the historian. What sort of evidence instructs the historian as to whether and how these analogies and truisms are to be applied ?

The " covering law " logicians do have an answer to this question as to the sort of evidence capable of supporting a claim that " it is entirely natural " that an event " happened as it did " because of such and such conditions. This evidence takes the form of correlations between kinds of properties, events, etc., which are " independently established "(p. 265). It is on the basis of statements in support of just this view that Gardiner is identified by Dray as holding the " covering law theory ". Dray is certainly correct in maintaining that proponents of this theory have never made explicit exactly what sorts of covering laws were appropriate and exactly what the logic of their employment is. Indeed, in his view, a characteristic carelessness about such details compromises the theory. But is it not rather the case that for supporters of the theory, these details are relatively secondary considerations, secondary in the sense that taking note of them would not affect the status of the essentials of the theory ? For example, this seems to be Ernest Nagel's position in the following passages from his paper, " Some Issues in the Logic of Historical Analysis ", where he is speaking of " the general pattern of explanations in historical inquiry ". He writes :

> The explanatory premises in history, as in the natural sciences, include a number of implicitly assumed laws, as well as many

540

exlicitly (though usually incompletely) formulated singular statements of initial conditions. The tacitly assumed laws may be of various kinds. They may be statements of regularities well attested in some special science, or they may be uncodified assumptions taken from common experience ; they may be universal statements of invariable concomitance, or they may be statistical in form ; they may assert a uniformity in temporal sequence, or they may assert some relation of co-existent dependence.

Doubtless the basic trouble in this area of inquiry is that we do not possess at present a generally accepted, explicitly formulated, and fully comprehensive scheme for weighing the evidence for any arbitrarily given hypothesis so that the logical worth of alternate conclusions relative to the evidence available for each can be compared . . . Fortunately, though the range of possible disagreement concerning the force of evidence for a given statement is theoretically limitless, there is substantial agreement among men experienced in relevant matters on the relative probabilities to be assigned to many hypotheses. Such agreement indicates that, despite the absence of an explicitly formulated logic, many unformulated habits of thought embody factually warrantable principles of inference.[1]

What is Dray's own answer to the question raised above ? As Strawson observes (p. 266), Dray encourages us to think " that generality has no place in particular historical explanation ". According to him, no independently established correlations need be tacitly assumed. How does the historian judge concerning causes ? The process is shrouded in mystery. Strawson, on the other hand, contends that analogies and truisms, general knowledge of human nature and public affairs, guide the historian's judgment. But even his very brief remarks are misleading and their inadequacies underscore those of Dray's account.

In terms of the " covering law theory ", as remarked above, the historian's explanation remains unsupported unless he can invoke independently established correlations between kinds of properties or combinations of properties. Strawson speaks of analogies and truisms as guiding the historian. But in approving Dray's critique of " covering law theory ", Strawson pointedly remarks that the historian's explanation could not be established " independently of the particular case concerned ". To think otherwise is " absurd " (p. 266). Here, indeed, is the crux of the matter. Just as Strawson asks Dray how the historian can judge without any recourse to generalizations, so one is inclined to ask Strawson how the historian can defend his use of these generalizations (the analogies, truisms or general knowledge of whatever sort) by reference to the particular case. Here too is a mystery.

I suggest taking the covering law stress on independently established correlations more seriously. Nagel's careful statement about " unformulated habits of thought " on the part of the historian has

[1] In P. Gardiner, *Theories of History* (Glencoe, Ll., 1959), pp. 376 and 385.

the virtue of locating these processes which resist logical formulation within a clearly formulated logical pattern. How the historian weighs the evidence remains unformulated. But we can at least say that the indispensable warrant for his judgment is independently established general correlations of some sort.

The failure to preserve this grain of truth in the " covering law theory " seems to be due to certain undeniable considerations that Dray properly emphasizes. The historian rarely, if ever, finds generalizations ready at hand under which he can directly subsume the event to be explained. (To explain the French Revolution is not to apply to it a theory or law of revolutions in general.) He cannot even readily subsume the event under a combination of such already verified generalizations. (Dray discusses this in connection with what he calls the complex version of the " covering law theory ".) Accordingly, when the historian's explanation is formulated as a general statement, such a statement is highly specific and it seems odd to think of such a generalization as independently established when it so obviously originated in an effort to explain one very particular event. Still perhaps an oddity is preferable to a mystery.

It is at least interesting to try to describe what is involved in historical judgments of this sort on the assumption that the only warrant for them takes the form of independently established general correlations. Let us imagine how one of those odd, highly specific generalizations comes about. The historian has to work out a general law covering the specifications of the event he is explaining. Every aspect of the event which he explains he explains by means of a judgment based on familiarity with the occurrence of such properties in other event-situations. Failing even one other totally similar event, he looks about for some events similar in some respects, other events similar in other respects. By means of a process of composite comparison he may reach a stage of inquiry at which he considers it warr nted to assert what some or all of the conditions were which made the occurrence of the event in question " natural ", even though all of the relevant properties of that event are not found in any of the events utilized in his comparisons.

The resultant explanation, that $C_1, C_2, C_3 \ldots C_n$ were the determining factors of the occurrence of E, is, in principle, supported by an exceedingly complex law, and, ideally, for every component of this law, the historian could cite supporting evidence in the form of significantly similar occurrences. Obviously, since he has to construct a law himself utilizing these data as evidence if he is to offer a warranted explanation in the first place, it will be necessary for him, in defending his conclusions, to produce the argument by which he concluded that the evidence supports his explanation. If he is unable to convince us that occurrences have taken place which do have evidential value for purposes of this inquiry, then his proposed explanation is deemed unwarranted. It is a candidate of the sort we were looking for when we asked " Why did E occur ? ", but a candidate without warranting credentials.

The complex supporting law which is at least implicit in the historian's reasoning, I shall call a covering law. It should not be confused with working generalizations concerning specific factors in the situation under scrutiny. These are often loose and porous, as Gardiner reminds us, and have to be refined in the course of historical investigation and explanation.

Dray insists upon " the peculiarities of the historian's typical problem—the weighing of a set of miscellaneous ' factors', which cannot be reduced to deduction from general laws . . .? ." [1] What is the weighing operation really ? It consists, I submit, in considering whether a certain circumstance was a necessary condition of the occurrence of an event. But particular circumstances or facts, viewed in splendid isolation from any comparable circumstances or facts, reveal nothing at all as to their weight. The expert historian, like the expert weight-guesser in the circus side-show, can arrive at a warranted conclusion concerning " weight " only by gauging according to general rules or laws. The same considerations apply to the utilization of analogies and truisms. If the historian is not sure which general knowledge to apply or how to do so, he must weigh the alternatives according to generalizations from past experience. The actual " weighing " is or involves deduction. To say that historical judgment as a process of weighing " cannot be reduced to deduction " can only mean that the historian has often to develop or refine the generalizations (by inductive techniques, however informal the process) which he uses in weighing. This latter procedure may be a peculiarity of the " historian's typical problem " (a questionable supposition) but does it involve a deviation from the logic of explanations as schematized by the covering law logicians ? Does it constitute as Dray indicates, reasoning of a different kind ? It would seem misleading to say so.

I would simply urge, in conclusion, that one need not make a fetish of the unity of science in order to appreciate the considerable merits of the covering law analysis of historical explanations. In stressing these merits I do not mean to pass a blanket negative judgment on Dray's book. It remains, in my opinion, an outstanding contemporary contribution to the philosophy of history.

ROBERT C. STOVER

Vassar College

[1] *Laws and Explanation in History* (London, 1957), p. 56.

I SHOULD like to reply to Mr. Scobie's comments (MIND, April 1960) on my discussion of the identity of a work of art (MIND, January 1959). Scobie complains there that " though we may use Margolis's division into ' types ', ' megatypes ', ' prime instances ' and ' prime notations ' when discussing the position of the artifact it can take us no farther than this, and can be no help at all in establishing the identity of a work of art, if by this we mean something other than the identity of the artifact. We have not even begun to distinguish the criteria required to establish in a certain few artifacts, their identity as works of art " (p. 258).

Now, I find this " objection " a little curious. In speaking of the identity of a work of art, I was, in my paper, clearly concerned with the question of how we individuate works of art and not with the question of the grounds on which we judge something to be or not to be a work of art. I recognize the second question to be an interesting one but a question altogether distinct from the one I was raising. And I think Mr. Scobie understood me to be discussing the first question at least, since he actually attempts to extend, in an interesting way, the sort of distinction I was exploring ; he does this, for instance, in distinguishing between sculpture casting made from a mould and from an original sculpture, prints made from lithograph stones and from woodblocks. I take it that, in pursuing such matters as these, Mr. Scobie is actually discussing how we individuate works of art of certain sorts by reference to relevant artifacts. I take it, that is, that Mr. Scobie is in implicit agreement with me on the very question I was hoping to illuminate. And I find, therefore, that his objection must be a reminder that other (and perhaps more important and more interesting) questions remain to be answered—in particular, that of the criteria by which we judge something to be a work of art.

But, on this score, I must insist that I was speaking (and Mr. Scobie in effect speaks) of the identity of a *work of art* (in the sense of providing for the individuation of works of art) and not of the identity of artifacts. How we individuate works of art poses something of a puzzle, but I took it that the individuation of artifacts did not, in our context, tax us in any way. So I cannot agree with Mr. Scobie's way of putting my account, that I was interested in " establishing . . . the identity of [an] artifact ". I suggest that he has confused the problem of the individuation of works of art with that of the problem of the definition of a work of art.

Several other of Scobie's remarks invite comment. For one, he holds that " the relationship of poem to translation, of the original musical manuscript (in other words, the known intentions of the composer) to performances in other keys and with instruments different from those stipulated in that score, of the play ' Hamlet ' to

544

the film of that name . . . have no bearing whatever upon . . . theories of art " (p. 256). I think he must be quite mistaken, though I see his point, particularly about translations ; at any rate, it is a matter quite indifferent to the issue of individuation and it is surely arguable—I should want to make much of the fact that music and drama are performing arts, and to argue that scores and scripts are, characteristically, incomplete notations, that variant performances and productions do not necessarily point to " unscrupulous producers . . . and actor-managers bent on self-glorification ", that we cannot say that deviations from the playwright's directions in the script " must be an error ", that one cannot draw ready conclusions for the performing arts from considerations applying to sculpture and painting (*cf.* pp. 257-258).

Secondly, Mr. Scobie seems to think that the question of what qualifies an object as a work of art is much the same as the question of what qualifies a man as an artist ; hence, his reference to the Van Meegeren forgeries and to copyists (pp. 256-257). I think the questions are distinct, but answerable by the same sort of effort (think for instance of a cathedral as a work of art and of the workmen not being, as such, artists).

Thirdly, Mr. Scobie seems to think that the question of what qualifies an object as a work of art is much the same as the question what sort of comments are relevant for the aesthetic appreciation of such an object ; hence, such comments as : " Translations of poetry have no aesthetic interest, except as literature in their own language, and then of course as something quite apart from the original poem " (p. 257). I should say that Scobie's particular point about translations is debatable (we speak, for instance, about " the poem in translation," which suggests that translations have some aesthetic interest *as translations* ; in fact, if they are taken merely as poems in their own language, translations would very obviously not be considered as translations). Also, I am reminded that T. S. Eliot made much of the difference between Shakespeare and Dante in terms of the translatability of their poetry—but perhaps Mr. Scobie would not wish to disqualify translations in this respect. Furthermore, the translation model does not seem to me to apply, as Mr. Scobie would apparently insist, to music ; he speaks for example, of Rosalyn Tureck's pianoforte performances of Bach's harpsicord pieces as an " attempted translation " (p. 258). I should have thought this way of speaking would have weakened the case against translation ; but it is surely a mistaken way of viewing musical performance. I should add, finally, that the issue of the aesthetic value of a translation is a quite distinct question from that of individuating works of art.

And fourthly, Mr. Scobie seems to think that the question of what qualifies an object as a work of art is much the same as the question of what an artist's intentions may have been (in some appropriate context) ; hence, he comments that deviation from script " must be an error ", that lithograph stone inked by " a person not the artist "

with " colours and textures not those intended by him " results in " an artistic rape " (p. 257). For one thing, I believe these items bear on the the question of the identity of a work of art (in the sense mentioned) and not on that of qualifying as a work of art. And for another, Mr. Scobie has not adequately considered the question of the relevance of the artist's intentions to a work before us, performed or completed under the circumstances indicated. I suspect Mr. Scobie finds the artist's intentions very easy to detect and regards them as peculiarly important in judging whether certain perform-ances or certain artifacts produced under complex conditions qualify as works of art. I, on the other hand, find the artist's intentions usually difficult to track down, frequently irrelevant in judging works of art, and appealed to only in very special circumstances—for in-stance, in judging, whether a performance of Strindberg is an " authentic " production.

JOSEPH MARGOLIS

University of Cincinnati

PRIMARY AND SECONDARY USES OF
MORAL LANGUAGE

To say of moral discourse that it is used primarily in such and such a way may mean any one or combination of several things. One could mean that this discourse is *most* often used in a certain way. He could also mean that this discourse cannot be understood unless what the user intends to accomplish with it is known. Thus, discovering the primary use of moral discourse might be thought of as being roughly the same thing as discovering the function of this discourse. Finding the primary use of moral discourse, in this sense, would then answer such questions as ' What is it used for? ' ' What is its purpose? ' or even ' To what end was it created? '

A third way of interpreting ' The primary use of moral discourse is . . . ' is in terms of that discourse's distinctive features. Finally a fourth is in terms of the rules of the discourse themselves. Here one asks, in effect, which use of this discourse has logical priority in some sense over the other uses. ' In some sense ' is appropriate, since within the purview of this interpretation there are several difficult-to-classify ways in which a person might say that one usage has priority over another.

Nowell-Smith refers to some of these when he speaks of lying, play-acting, story telling and irony as being secondary uses, that is, ' . . . uses to which an expression could not (logically) be put unless it had some primary use ' (p. 81). There are other cases of secondary usage which he talks about, such as a descriptive use of ' good ' which can be used to mean ' . . . what most people call " good " ' (p. 166). He also discusses the use of criteria in moral contexts, arguing for their secondary status relative to our use of moral language when we have pro-attitudes toward something (p. 173). Still other cases fall within the framework of this fourth interpretation. Consider the case of A who has recently had a moral argument with his son a. In reporting this argument to B, who is in no way involved in the issue, A might say (and in effect, report his persuasive use of moral language) ' I told him " He ought to do it " because he ought to do it '. Apart from whether such a usage can be allowed within moral discourse, if it can be, then the latter non-persuasive use of moral discourse would be primary over the persuasive since it was used to *justify* the persuasive.

Now all these cases, and the list of examples need not be thought of as exhaustive, I take to have one thing in common ; *viz.* an appeal to the purely *internal* considerations of a discourse in making the primary-secondary use distinction. So even if the distinction can be made in ways which I have not mentioned, provided it is made by an appeal to the rules of the discourse alone, then it will qualify as falling under this fourth interpretation.

Although hardly complete, such a characterization of this interpretation of such phrases as ' the primary use of ' is nonetheless

adequate for distinguishing it from the other three. It should be
clear, for instance, that there is no reason why the primary use in
this so called *internal* sense must be used more frequently than some
secondary use. Nor need the primary use, again in the internal
sense, be the use which answers the question ' What is this discourse
used for? ' Nor, finally, need the use which distinguishes moral
discourse from any other discourse be identical with the primary
use in the internal sense.

A mistake then that analysts might make and, I suspect, have been
making, is to look for the distinctive use of a discourse and then say
of any other use within that discourse that it is secondary to the
distinctive use. Of course, labelling the distinctive use ' primary '
and the other uses ' secondary ' is in itself harmless, provided the
analyst knows what he is doing. But if such an analyst is confusing
his use of ' the primary and secondary uses ' with another use of this
phrase and, for example, says things which show him to be actually
commenting about the internal rules of a discourse instead, it is
then that the trouble begins.

By way of strengthening the suspicion that this mistake is a com-
mon and dangerous one, one should first note the presence of such
assertions as ' The function of moral discourse is . . . ' at the
beginning of a work on ethics. Partly, of course, such assertions
are made at the beginning rather than at the end of a book for
stylistic reasons. But partly, I think, these statements are thought
to be obviously true by the author and this is why they are there.[1]
Now if these statements were intended as comments strictly speaking
about the rules of use, I can hardly imagine that they would be
thought to be obviously true for two reasons. In the first place,
what usage is primary or secondary could only be established *after*
the analyst knows the whole " game ". Until then, he could at
best say of one usage only that it is secondary relative to another, but
not that the latter is primary. Secondly, as has already been
pointed out, the primary usage may be infrequently or even rarely
employed and, thus, far from being something we all know about,
it may be discovered only *after* careful analysis.

The second thing to note about the alleged mistake in recent works
on ethics is how these " obvious " comments about ethical discourse
are used to solve the traditional ethical problems. This is not to
imply that other, legitimate, criticisms are not levelled at, say, the
cognitivist. Rather it is to argue that once these writers have found

---

[1] After saying, " The central activities for which moral language is used
are choosing and advising others to choose " (11), Nowell-Smith comments
on the very next page, " But the great philosophers of the past always
treated questions of appraisal as subordinate to practical questions.
They assumed—*and who would not?*—that the point of telling you that
Jones is a good or a bad man is that you should imitate or not imitate Jones,
that you should or should not give Jones the job or whatever else might
be in question " (italics mine).

moral discourse to be *obviously* used primarily (in any except the internal or narrow sense) for achieving some practical end, they leave their reader with the impression that this discovery should upset the cognitivist. This discovery purportedly either helps to show how he has gone wrong [1] or it shows this cognitivist to be dealing with what are not really the important things in ethics.[2]

I wish to contend against both of these alternatives. Against the first it must be admitted that if an analyst's interest so inclines him, he could inquire into the question of how moral differs from empirical discourse. He might very well then discover that the former discourse is practical in a way in which the latter is not. He must, however, be aware of what his interest will yield. Should he expect from it some practical consequences, such as helping the ordinary man deal with moral concepts, all well and good. If, instead, he expects such an analysis to yield answers to the issue separating the cognitivist and non-cognitivist, then he is mistaken. His kind of analysis will not even by-pass such a question. The kind of analysis which will, concentrates on the " grammar " of moral discourse in the sense of analysing the rules *internal* to the discourse. It is true that in handling the question of cognitivism or, for that matter, the issue of the reductionist versus the non-reductionist, one must look to other discourses to find such things as the paradigm case of a cognitivist discourse. Nevertheless, once the use(s) of ' cognitive discourse ' has been established, ethical discourse turns out to be either cognitivist or non-cognitivist as a result of an analysis of its " grammar " in the narrow or internal sense of the word. To say then that moral discourse is practical because its aim, purpose or distinctive features involves guiding behaviour is not as yet to exclude the possibility that it is also cognitivist in some other sense. Putting this another way, analyses of the practical uses of moral discourse do not, finally, show how the cognitivist has gone wrong.

---

[1] The following passages suggest, at least, the conflict Nowell-Smith feels exists between the practical and the " theoretical " roles of a discourse: (1) " But it might be objected that the fundamental question is not ' What shall I do? ' but ' What ought I to do? ' and the fundamental concept not decision but obligation. My reason for treating the ' shall ' question as fundamental is that moral discourse is practical. The language of ' ought ' is intelligible only in the context of practical questions, and we have not answered a practical question until we have reached a decision. ' I shall do this ' is the general formula for expressing decision " (p. 267) ; (2) " We ask for help in the solution of practical problems and they offer us a description if a non-natural world " (p. 61).

[2] " Discussing these doctrines (objectivism, subjectivism, and imperativism—my add) in turn will be enough of a programme to start with. If we are to be satisfied with any of the traditional theories, it must be such as will help us to tell good reasoning from bad—we therefore have a test to apply in criticizing them. And afterwards, if none of them helps us to find the answer to our central, practical question, we can always return and attack the problem head-on " (Toulmin, p. 5).

The way a non-cognitivist must refute the cognitivist, and the only way to do it, is for the non-cognitivist to demonstrate that there is no unique referring (or asserting) use present in moral discourse.

Against the second alternative, I think it can *now* be shown that the so-called cognitivist *is* dealing with an important issue in ethics. For understanding the point of the variety of uses to which moral language can be put, the cognitivist can hardly avoid asking himself which type of use, if any, deserves to be called primary. Clearly his answer must be that the alleged uniquely-referring (or asserting) use fits the bill as *the* primary use. In other words, the cognitivist is not, as is commonly supposed, merely arguing for the presence of a unique cognitive feature in moral discourse. He is insisting, further, that an analysis of the rules " internal " to this discourse will reveal all other uses to be parasitic upon the cognitive one. Even more, this cognitive use serves to *justify* the other uses (see p. 2). Without it, moral discourse would be something quite other than what it is for the cognitivist.

It is most misleading then, to accuse the cognitivist of dealing with questions which are not important. It may be that the cognitivist's answers to these questions are wrong, but this is another matter. But even here part of the burden of this note has been to show that the activities of most analysts have not effectively challenged the cognitivist's claims. These analysts have been doing interesting things such as pointing to the variety of practical uses of moral discourse. But, by and large, they have not met the cognitivist head on, probably because they have failed to understand what he is saying. They have failed to understand the implications of his position with respect to the various possible uses of ' The primary use of moral discourse is . . . .'.

N. FOTION

Coe College

In his book *An Analysis of Knowing* Mr. Hartland-Swann presents a decision-analysis of knowing *that* statements. There seems to me some ambiguity in his analysis of the decision element involved in a sense-statement. He says, " Suppose . . . someone says to me ' I have a pain in my knee-cap ' and that, facetiously perhaps, I say ' Are you stating a fact ? ' The expected reply—' of course I am '— can only be made because, owing to certain neural reactions, the person concerned has decided that he *has* a pain in one particular place, namely, his knee-cap " (p. 21). In this passage the statement, owing to certain neural reactions the person concerned has decided that he has a pain in his knee-cap, is very ambiguous. A similar statement may be constructed in connection with a colour-statement, such as, ' I am sensing a red patch of colour '. Following Hartland-Swann one may say, owing to a certain wave-length the person concerned has decided that he is sensing a red patch of colour. The former statement seems to me as absurd as the latter. Hartland-Swann may mean by his remarks that the person who says ' of course I am ', is able to give such a reply because certain neural reactions have *caused* in him the decision that he has a pain in his knee-cap. But such an interpretation makes no sense. For, certain neural reactions may cause a pain in my knee-cap or in any other part of my body, but it makes no sense to say that my decision that I have a pain in my knee-cap has been caused by these neural reactions. In the same way a certain wave-length may cause a red patch of colour but it makes no sense to say that my decision that I am sensing a red patch of colour is caused by such a wave-length. The logic of the expression ' Having a pain ' is different from the logic of the expression ' Having decided that one has a pain '. A decision is not a thing to be caused as a pain in one's knee-cap is caused by certain neural reactions. Then what else is the meaning of Hartland-Swann's statement ? He may mean that certain neural reactions are the *basis* and not the *cause* of my decision that I have a pain. I have decided that I have a pain in my knee-cap because I happened to accept the decision that certain neural reactions cause a pain in one's knee-cap. My former decision is based on my latter decision. This implies that Hartland-Swann thinks that one has decided that one has a pain because one is in possession of a technical decision made by the neurologists. Suppose one does not know about this technical decision, in such a case would Hartland-Swann maintain that the person concerned cannot make the decision that he has a pain in his knee-cap ? Hartland-Swann may deny this conclusion. His words " The expected reply . . . can *only* be made " may mean that only in a particular case when one gives the reply, ' of course I am stating a fact ', one's decision that one has a pain in one's knee-cap is based on the technical decision. There may be cases in which one may not know at all that certain neural reactions

cause a pain in one's knee-cap, yet one may decide that one has a pain in one's knee-cap. However, if I decide that I have a pain on some non-technical ground, Hartland-Swann may mean, then I do not have a right to reply ' of course I am stating a fact '. Such a right can be given to me only when my decision is based on the technical decision as pointed out above. Even this qualified claim is illegitimate. For, no sense-statement can be based on a technical decision. This follows from Hartland-Swann's own view. He makes a distinction in the meaning of the expressions ' in accordance with a dominant decision ' and ' in accordance with a personal decision '. He thinks that all empirical statements are directly or indirectly in accordance with a dominant decision. The same is not true about sense-statements. These have been claimed by Hartland-Swann to be in accordance with merely one's personal decision. Once it is granted that one's decision that one has a pain is based on a technical decision, it is granted that one's sense-statement is indirectly in accordance with a dominant decision. A technical decision, even if only one man holds it, is capable of becoming a dominant decision. So if one believes that a sense-statement is in accordance with a technical decision one has also to believe that it is in accordance with a dominant decision. This shows that Hartland-Swann cannot allow the decision concerning a sense-statement to be based on a decision which is technical. So it is not only the case that certain neural reactions cannot be taken as the cause of my decision that I have a pain in my knee-cap, they can also not be taken as the basis of such a decision. I wonder why Hartland-Swann has introduced these ' neural reactions ' in his decision-analysis of sense-statements ? There seems to me no other reason than that he wants to kill a tiger with the bullet for a bird. An analysis which may run smoothly in connection with such statements as ' Gandhi is dead ' and ' My daughter will marry you ' may not run so smoothly in connection with ' I have a pain '. Hartland-Swann wants to put all sorts of statements into one pigeon-hole, so he is committed to make ambiguous statements.

SURESH

Indian Institute of Philosophy,
  Amalner, E. Khandesh, Bombay

Causation in the Law. By H. L. A. HART and A. M. HONORÉ.
Oxford : The Clarendon Press, 1959. Pp. xxxii + 454. 55s.

THE intention of this book is primarily to be of help to lawyers, not to philosophers. The authors' first aim is to " identify the sources of the uncertainties and confusions which continue to surround the legal use of causal language in spite of a vast juristic literature dedicated to its clarification ". Since they believe that the law's use of causal language embodies, though it also extends, principles inherent in the plain man's concept of causation, that concept is examined in Part I. This part is likely to be of most interest to philosophers.

Before turning to legal matters, the authors give us some penetrating criticisms of the Hume-Mill tradition, a clear and often brilliant mapping of the whole area of causal language, with special emphasis on the distinction between a cause and a mere condition, an account of the ways in which generalizations are applied to particular cases and an account of the contrast between causes and reasons. As we should expect, there is much emphasis on the variety and complexity of causal language—the account of the difference between consequences, results and effects (24-26) is particularly good—, the variety and complexity is shown to be neither arbitrary nor unnecessarily elaborate, and the confusions and uncertainties are traced mainly to its neglect. The general thesis, with its emphasis on the part played by the concept of causation in inquiries that are particular and practical, is Collingwoodian ; the influence, though not the style of Wittgenstein and Austin is everywhere apparent. This part can hardly be too highly praised ; had the authors' interests been directly philosophical it might have been longer ; for the treatment of some topics (voluntary action, for example) does not pretend to do more than is required for the exposition in Part II.

It does not take long to show that the traditional attempt to analyse causation in terms of necessary and sufficient conditions is inadequate and largely irrelevant to the task of throwing light on those causal problems that arise in law and in ordinary practical life. It is not so much that Hume was wrong as that he simply has nothing to say about these problems. He was concerned with the metaphysical foundations of natural science and with the discovery of general laws, not with the explanation of particular events that require explanation precisely because they are unusual, still less with the allocation of responsibility. There is no room in his system for the crucial distinction between a cause and a mere condition which looms larger as the book proceeds. Mill is found to be more helpful since he recognized this distinction, without however getting to the bottom of it, and since he loosened up the whole discussion (12-23).

The central conception of causation (26-30), from which others are developed by various analogies, is that of the purposive manipulation of objects by people.

> Human beings have learnt, by making appropriate movements of their bodies, to bring about desired alterations in objects, animate or inanimate, in their environment, and to express these simple achievements by transitive verbs like push, pull, bend, twist, break, injure. . . . Men have, however, learnt to extend the range of their actions and have discovered that by doing these relatively simple actions they can, in favourable circumstances, bring about secondary changes, not only in the objects actually manipulated, but in other objects. . . . Here we use the correlative terms ' cause' and ' effect ' rather than simple transitive verbs : the effect is the desired secondary change and the cause is our action in bringing about the primary change in the things manipulated or those primary changes themselves. . . . Here the notions of cause and effect come together with the notion of means to ends and of producing one thing by doing another.

The primary notion of cause is, then, that of an action which ' makes a difference ' and this notion clearly presupposes that of a regular order or course of events which would take place but for the action called a cause. The answer to the question that bothered Hume is thus quietly assumed.

This central notion of cause is developed in three main ways. (i) *Natural events* which are thought of as ' making a difference ' come to called causes by analogy with human actions that make a difference. (ii) Since men often bring about desired results by talking to other men and since the manner in which this is done differs in striking ways from the manner in which they bring about changes in remote objects, we get the concept of an *interpersonal transaction* (48-57), a situation in which one man gets another to do something by means of threats, bribes, inducements, commands, advice or statements, true or false. While the authors are no friends to *L'Homme Machine* the discussion here should put an end to too glib a use of the contrast between causes and reasons. There are too many mixed and borderline cases. (iii) " Men have discovered that nature is not only sometimes harmful *if* we intervene, but is also sometimes harmful *unless* we intervene, and have developed customary techniques, procedures and routines to counteract such harm " (35).

We lock our doors, fence in animals and give conventional warnings of danger ; and these precautions have come to be as much a part of the regular background against which human action takes place as are the natural regularities themselves. Consequently an *omission* to take a common precaution, since it is an interruption of the normal, though (in one sense) not natural background of activity can rank as a cause.

This is perhaps a good point at which to break off in order to illustrate the trouble that can arise when an important and pervasive

element in the notion of cause is mistakenly thought to be an essential element. The plain man has no hesitation in singling out someone's omission to take a common precaution as the cause of some subsequent harm; but the less plain lawyer may puzzle himself as to how something so inactive, so devoid of force as an omission can be a cause at all (28). In the mind of the judge or jurist there is often a tension between the plain man who knows that X was the cause of Y and the theorist who feels that before he can call it a cause he must find some way of attributing active potency to it. The attempt to do this sometimes issues in the most barbarous and fantastic language.

Causes and Conditions. Here the most obvious distinction is that between an act or other disturbance of the *status quo* and the ' standing conditions ' but for which it would not take effect: striking a match is a cause, presence of oxygen a condition. But this distinction does not cut deep enough; for, although the presence of oxygen would not usually be a cause, it might be. " We have only to consider a factory where delicate manufacturing processes are carried on, requiring the exclusion of oxygen, to make it perfectly sensible to identify as the cause of a fire the presence of oxygen introduced by someone's mistake " (10).

When someone has been deliberately shot, why is it absurd to cite as the cause of death either the fact that his blood cells have been deprived of oxygen or the sale of the gun to the murderer's father from whom he inherited it? Why and how do we pick on the murderer's act to be the cause? These are connected questions, since both the selection and the principles on which it is made will depend on the purpose of the inquiry. Why determines both what and how. The crucial question here is never ' what generalization can be invoked or discovered to explain this phenomenon? ', but always ' in what ways can familiar generalizations be combined and applied to this case? '; and this question gives rise to a range of difficulties with which philosophers have not much concerned themselves. This is surprising since, as the authors abundantly show, much of the detail of our causal language reflects our attributive rather than our explanatory interests.

Breaking the chain of causation. " Typically what precipitates these difficulties is that, among the conditions required to account for the harm which has occurred, there is found in addition to the defendant's action a factor (usually a human action or some striking natural phenomenon) which itself has some of the characteristics by which common sense distinguishes causes from mere conditions; so that there seems as much reason to attribute the harm to this third factor as to the defendant's action" (23). This leads to the type of problem that figures most prominently in the book and so, we may judge, in the law. The length of Cleopatra's nose was not, it is clear, the cause of the fall of the Roman Empire. When and how can the causal efficacy of an action be ' traced through '

subsequent events to some harm that remains its consequence? When does a third factor, intervening between necessary condition and subsequent harm negative causal connexion or " in the terms of the law's favourite though perhaps most misleading metaphor, break the chain of causation? "

Two preliminary but important points : (*a*) Though we now have more than one candidate for the role of cause, our selection is not arbitrary. The law is entitled to cut off *liability* where it likes—that is another matter of which more later—but it cannot arbitrarily decide where *causal connexion* is cut off. That is determined by the application to the facts of principles embodied in our ordinary causal concepts. (*b*) The law is to some extent compelled to extend and distort causal language, because the courts labour under two disadvantages from which the plain man is free. (i) The plain man can simply refuse to give any decision at all ; his stock of causal principles gives no unambiguous answer in the case ; or he can say that it was partly caused by one thing and partly by another. The courts, however, *must* decide, and, for historical reasons, they make less use of apportionment than the plain man would think sensible and just. (ii) The plain man can select from his large repertory of causal expressions the one that seems best to fit the case, and will often avoid the *word* ' cause '. " A's blow made B stagger and break the glass ", he will say, rather than ' A caused the glass to break '; but the law has to decide whether A's act was the proximate cause of the damage in those very terms. (This may partly be due to an historical consideration that the authors do not mention. Though the old ' forms of action ', set formulae in terms of which pleas had to be made out, were abolished a century ago, their spirit still remains ; both plaintiff and defendant must make out their pleas in terms which the law recognizes ; for only so can they be brought under the principles and precedents which guide the court. One consequence of this restriction of vocabulary is important and is recognized by the authors. It is partly because of the strained and artificial uses to which the *word* ' cause ' is put that the modernists, whom we shall meet in a moment, have to come to disparage the *concept*.)

To return, however, to the chain of causation : according to the authors, causal connexion between a necessary condition and a subsequent event is cut off in two main ways : (*a*) by the intervention of some very abnormal circumstance, or (*b*) by a *fully* voluntary human act. (*a*) If a man negligently starts a fire and an evening breeze springs up so that it spreads to a forest, the intervening breeze, being a normal thing, does not negative causal connexion between the negligent act and the harm. But if the fire would not have spread to the forest but for the intervention of a most unusual hurricane, the negligent act can no longer be cited as the cause. (*b*) A man negligently leaves a hole in the road into which another accidentally pushes his friend so that he breaks his leg. The

negligent act is the cause of the harm and the negligent agent will be liable. But if an escaping prisoner pushes the sheriff into the hole so that he breaks his leg, the initial negligence is no longer the cause of the damage. It is important to note that, to break the chain of causation, a human act must be *fully* voluntary and that there is a long list of exceptions through which causality can still be traced (134-151). Besides such obvious cases as shock, accident and mistake, the following types of action do not count as fully voluntary : actions done in self-preservation, preservation of property, safeguarding of rights and interests, fulfilling legal or moral obligations. This is presumably because the act is, in a way, forced on the agent by the predicament in which he is placed.

It is at first sight puzzling to see why abnormal events and voluntary acts should be thought to *break* the chain of causation when they so obviously *make* it in the sense that they are necessary links in the chain between initial cause and ultimate effect. The authors' explanation is so brilliant and so aptly illustrates their method that it must be quoted in full.

> When we understand the power exerted over our ordinary thought by the conception that causing harm is a mere extension of the primary case of doing harm, the inter-related metaphors which seem natural to lawyers and laymen, in describing various aspects of causal connexion, fall into place and we can discuss their factual basis. The persistent notion that some kinds of event required in addition to the initiating action for the production of harm ' break the chain of causation ' is intelligible, if we remember that, though such events actually complete the *explanation* of the harm (and so *make* rather than *break* the causal explanation) they do, unlike mere normal conditions, break the *analogy* with cases of simple actions. The same analogy accounts for the description of these factors as ' new actions ' ' new causes ', ' superseding ', ' extraneous ', ' intervening forces ' : and for the description of the initiating action, when the chain of causation is broken, as ' no longer operative ', having ' worn out ', ' *functus officio* ' (69).

Illuminating though this is, it leaves some doubts unresolved. As philosophers know all too well it is difficult to draw the line between those acts that are fully voluntary and those that are not. Was not the escaping prisoner under some pressure to push the pursuing sheriff into the hole? Nor is it easy to see just *why* factors of these two kinds should be held to negative causal connexion while others do not. In the case of abnormal events, the idea seems to be something like this. When I let in the clutch the causal efficacy of this simple movement can be traced through the intervening mechanism to the movement of the car, not only because this mechanism is normal and regular, but, more importantly, because I make use of this known regularity, so that ' setting the car in motion ' is what I intend to do. Everything from movement of foot to movement of car constitutes one action. But obviously I cannot intend to do X by performing the simple action Y if an abnormal event has to intervene between Y and X.

This explanation, however, will not account for a doctrine according to which normal events, *not* known to the agent and hence not exploited by him, fail to break the chain. In general it will not explain why the consequences of *negligent* acts can be traced through normal events.

Finally it is not clear whether or not the two different types of third factor which break the chain have a common basis. If they do, the idea must presumably be that there is something abnormal about a fully voluntary act. But this is just not so. If A is induced, bribed or threatened by B into doing some harm, B may be liable on the grounds that A's act was not fully voluntary ; but he is not liable if he merely advises A. Yet it is no more *abnormal* to take advice than to take a bribe. The authors are no doubt right in their account of the conceptual facts ; in the first set of cases we should in fact say that B made A do it, but not in the last. But when we come to ask *why* this should be so, it seems that the old (and very dubious) idea that the fully voluntary acts of others are unpredictable must be at work.

II

Part II, from the authors' point of view the most important, is mainly concerned with a dispute among jurists which may need to be explained to readers of MIND. We must distinguish first between the cutting off of *liability* and the cutting off of *causal connexion*. A New York statute lays down that a man who negligently starts a fire is liable to pay only for the first of several houses it destroys (63). This is a limit on liability imposed for reasons of legal or social policy unconnected with causal principles ; for there is nothing in the nature of the world or in our use of causal concepts which leads us to say that a dropped cigarette can cause the destruction of one house but not of more. But, though liability and causal connexion are different, the one being ' arbitrary ' the other ' factual ', the one often follows the other. In enormously many cases a negligent agent is liable *because* his negligent act is held to have caused the harm.

At this point we come upon the grand terminological question that made the whole of Part I necessary. In what sense must the negligent act have been a cause? (I confine myself to negligence ; but the same question arises in other branches of the law.) This is where the dispute begins. According to the traditional doctrine, which on the whole the authors support, it must have been a cause in some sense much stronger than that of necessary condition or *sine qua non*. The courts take seriously the question whether the act was a cause in this stronger sense and to determine it they apply principles drawn from our common concept of cause elucidated in Part I. According to the ' modernists ', on the other hand, the purely factual part of the inquiry is limited to deciding that

the defendant's act was a necessary condition, usually a fairly easy matter. Once it is settled that the harm would not have occurred but for the defendant's act we pass *straight* to the question of liability, and this is to be settled on grounds of legal and social policy. Roughly, we ask whether it would be a good thing to hold defendant liable in such circumstances. Attempts to discover whether the act was a cause in a stronger sense are a mere blind ; they lead us into hopeless complexities and absurdities, and they cover up, with their misleading talk of proximate causes, chains of causation, remoteness of liability, etc., the principles of policy that really determine the decision.

According to the modernists, instead of pursuing further futile causal inquiries, the courts should address themselves to either or both the following questions : (*a*) what was defendant's degree of moral guilt, in comparison, perhaps, with that of plaintiff or a third party? (*b*) was the harm of which defendant's act was a necessary condition foreseeable or ' within the risk '? To explain this last phrase, a negligent act in law is not any act done in an off-hand way ; it is one likely to do some harm that a reasonable and prudent man would have foreseen and guarded against. While allowing both that there is an increasing tendency to apply the ' risk ' theory and that the modernist approach introduces much clarification, the authors decide against it for reasons that can be thus summarized.

(*a*) It will not do as an account of the actual decisions of the courts or of the reasons given in support of them ; for causal inquiries in the law are not as spurious as the modernists make out.

(*b*) It will not do as a recommendation to the courts to scrap causal inquiries in favour of asking the other questions proposed. The risk theory has something to be said for it when it takes the form of *extending* liability to cover cases in which, on a strictly causal view, it would be cut off. But it is not a good way of *limiting* liability in cases in which, on a causal theory, defendant would be liable ; for it has neither the consistency, the simplicity nor the fairness claimed for it.

(*c*) Apart from the arguments brought against it from the point of view of producing unfair results, there is a strictly logical argument against the risk theory unless it is prepared to take the heroic step of denying all liability for " ulterior harm". It is shown (238-243) that the risk theory covertly re-introduces causal questions with which it cannot dispense. " Indeed, what is true of nature is true of causation : *expellas furca ; tamen usque recurret* " (124).

To criticize the authors' defence of causation against the partisans of the risk theory requires more space and more legal knowledge than I have. The main argument seems cogent and the whole treatment, involving as it does some eight hundred cases many of which display features of an improbability that the wildest philo-sophical imagination would not match, provides a standing warning

against glib theorizing in this treacherous and tangled terrain. Simple generalizations that look promising too often founder on some case that has been decided otherwise and, we feel, rightly decided. There can be no doubt that lawyers will now have a truer picture of this part of their subject than has usually been painted. Yet a large philosophical doubt remains.

The general form of a legal argument is a simple syllogism with a factual minor premise ' A did X in conditions C ' and a legal (? synthetic *a priori*) major, ' Anyone who does X in conditions C is liable to legal consequences (penalty or damages) S '. The issue between traditionalists and modernists can then be seen as concerned with the question whether causal inquiries which extend beyond establishing that the act in question was a necessary condition of the subsequent harm are designed to establish the factual minor or the legal major ; and, seen in this way, the problem is one of a kind that has many other examples in philosophy. The authors claim that many such inquiries are genuinely factual and that the principles of law advocated by the modernists have causal considerations smuggled into them. But are not principles of policy smuggled in much the same way into the authors' principles of causality? Is not what Horace says of nature and the authors say of causality true also of legal policy? Though we try to expel it from the purely causal part of the inquiry, it comes obstinately back. The basis of these doubts can only be briefly sketched.

(i) The principles by which we are said to distinguish a cause from a mere condition and to decide whether causal connexion can be traced through or is cut off are what they are because we use causal language in an *attributive* way. To attribute is neither to describe nor to explain ; what, then, is it? It is tempting to answer that it is to fix on the agent the responsibility for the subsequent harm, and this comes perilously near to saying that the concept of cause, when it goes beyond that of necessary condition, presupposes the concept of liability. The authors recognize (61) the dual role of the expression ' he is responsible for it ', meaning both ' he is liable (to pay) for it ' and ' he did it (caused the harm) ' ; but to say as they do that the connexion between these two cannot be analytic since causality is here the *ground* for liability is too brisk. Causality is not as factual as it seems if the idea of imputing responsibility must be used for establishing it.

(ii) This consideration is re-inforced by the fact that in attributive contexts our decision as to whether something is or is not a cause depends on the *purpose* of the inquiry. Now since the purpose of legal inquiries is to assign liability, if this dependence is *logical*, considerations of legal policy are built into causal principles from the start.

(iii) The two crucial distinctions used for deciding whether causal connexion is cut off (normal/abnormal and voluntary/nonvoluntary) are not as purely factual as they seem. Nature knows nothing of abnormality ; it is we who decide what is sufficiently

abnormal to break causal connexion. *How* abnormal must it be?
And why should just *this* degree of abnormality be held to cut off
causal connexion when a lesser degree would not? It is difficult
not to answer that an event is sufficiently abnormal to cut off
causal connexion if and only if it is sufficiently abnormal to cut off
the agent's responsibility. For it is easy to see why the occurrence
of an event so abnormal that a reasonable and prudent man could
not be expected to take it into his calculations should excuse even a
man who has not been reasonable and prudent ; but it is not easy
to see why such an event should cut off *causal connexion* except *via*
this route which goes through cutting off responsibility. The similar
difficulty that arises over establishing whether an action is fully
voluntary without covertly introducing the idea of being responsible
for the action is too notorious to need comment.

(iv) The twin ' third factors ' that cut off causality have a
suspicious correspondence to the twin policy-substitutes of the
modernists, foreseeability and moral fault. For, by and large,
those consequences that are too remote for causal connexion on the
grounds of the occurrence of an abnormal event will be just those
that are not foreseeable by the agent or ' within the risk ' that
makes his act negligent ; and the idea that the fully voluntary act
of another breaks the causal chain looks as if it may depend on
the idea that there is now another agent on whom liability should
fall. We should not, of course, want it to fall on a wholly innocent
plaintiff or third party ; but that can usually be avoided by showing
that such acts come under one of the headings that make acts not
fully voluntary, for example by saying that the new agent, though
acting deliberately, was ignorant of the situation created by the
original tort. We are told that " the general principle of the tradi-
tional doctrine is that the free, deliberate and informed act or
omission of a human being, *intended to produce the consequence that
is in fact produced,* negatives causal connexion " (129 ; my italics).
I can find no case in which the wholly *innocent* voluntary action of
plaintiff or a third party is allowed to negative causal connexion
between defendant's negligent act and the subsequent harm. This
is striking, since it shows that the voluntary character of an act
does not, as such, break causation ; it does so only when it can be
used as an argument for shifting liability onto someone else.

It is tantalizing to guess what the authors' reply to this heretical
suggestion would be. They show that, over a very large area, the
application of causal and non-causal tests gives the same results—
a fact which in itself suggests that foreseeability and moral fault are
built into the concept of causal connexion—but they nowhere give
a *thorough* account of a case in which the two approaches give
different results with a full explanation of exactly why they do so.
This would, perhaps, have been of little interest to lawyers   but it
would have been of much to philosophers.

P. NOWELL-SMITH

University of Leicester

X.—NEW BOOKS

Humanism and Moral Theory. By R. Osborn. London : George Allen and Unwin, 1959. Pp. 115. 18s.

Although subtitled ' a psychological and social enquiry ', this book would make a good introduction to moral philosophy. The author believes, as most beginners would, that there is some considerable connection between moral views and empirical facts, so that the reader is not put off, as is sometimes the case, by a sharp distinction between facts and values that he is sure, in the early stages, to misunderstand. The tone is thoughtful, lucid and urbane and, although there is nothing startlingly new, only those who think seriousness cannot exist without anxiety are likely to find the content trivial. Mr. Osborn will not, in all eyes, prove all his main points ; in this short book he never pauses to cope with the relentless philosopher who would press critical enquiry to the limit. But this does not mean that his thesis is superficial.

Osborn starts from the obvious and obviously puzzling fact that people can agree very largely about what it is right or wrong to do yet disagree point blank about what it is to say that an act is right or wrong. (Captious critics will condemn his terminology here ; for it is out of line with current philososophical practice ; but the point is made clearly enough.) This leads him to an examination of the ways in which we learn to think morally in childhood and in which diverse experiences can produce diversity not so much in our moral views as in the spirit in which we hold them. The prime divergence here is that between a deontological and a teleological stucturing of moral thought, a divergence that exists not only among philosophers but, very obviously, in all ranks of society. There are some whose characteristic form of moral question is ' is it done ? ' and others whose characteristic form is ' what good will it do ? '. Osborn's great merit here is that he tells us where to look in order to solve this puzzle.

He takes the Aristotelian line that man is fundamentally a social and rational animal, but only potentially so. There is some evidence, not scientifically inexpugnable but strong enough, that a human baby not subjected to a human environment, in particular to contact with language-users, never grows up into a being actually social and rational. Starting from here, Osborn draws on Piaget and Freud to explain just how the potency is transformed into actuality. We may doubt the empirical findings, (though, in the case of Piaget, I should not) but it is up to anyone who doubts them to meet Piaget on his own ground and to show where his scientific reasoning has gone wrong. Armchair philosophizing will not do at *this* point. What Piaget shows is that a heteronomous, deontological stage of thinking about rules always precedes an autonomous stage. Roughly, children from five to nine think of rules as eternal, sacred, always to be obeyed and never to be questioned, because they are ' there ' as much as sunrise and sunset. Older children ask and try to answer questions about the *point* of a rule, they see rules as recipes for the achievement of ends, to be adopted, adapted or discarded entirely in the light of their empirically verified tendency to promote those ends. These empirical studies only comfirm (and the confirmation is not unnecessary) what we should expect. A child must learn *that* something is wrong or forbidden, long before he can begin to ask *why* it is wrong or forbidden ; so that his moral thought must be deontological before it can be teleological. This is not a ' must ' of logical necessity ; it is logically possible for a moral Pallas Athene

to spring from the head of Zeus. But given what we all know about babies and young children, it *does* follow. In Osborn's eyes a deontological structuring of morality is simply a case of development arrested at a pre-rational stage; and this by itself goes far to explain the original puzzle. Almost all the familiar moral rules have, or are believed to have, the consequence of increasing the rationality and sociability of the human animal; so it is not surprising that people who think rules binding as rules and people who treat them as recipes for achieving a certain kind of life manage to agree, by and large, on their content.

Osborn has no difficulty, then, in showing that a flexible, teleological outlook is sociologically more mature; but is not this to say that it is therefore *better* simply to commit the naturalistic fallacy? "If children naturally tend to transcend their egocentric interests, to think rationally and universally in terms of othe. as well as themselves, we can see the uninterrupted achievement of this state of psychological maturity as an ideal of human development—as the good for man. The ideal gives us a standpoint from which to assess the moral quality of different societies. One society is better then another to the extent that it provides the milieu in which men can realize their rational humanity." Phrases like ' rational humanity ' would require a much larger book to give them the filling out they need; but this is not an example of the fallacy of deducing value judgments from empirical facts. For the words "we *can* see" are an invitation, not a logical threat. Readers may refuse to see things in this way, may claim that the typically childish standpoint is *better* than the typically adult, if they wish. But if they are adult, they will probably not find this a comfortable position.

In one respect Osborn sticks too closely to his Greek masters. His declared aim is to "look for an interpretation of ethics in humanist terms, an interpretation which does not go beyond the bounds of human life for its criteria of good and bad, right and wrong, yet avoids the difficulties of relativism that the rejection of absolute, extra-human standards seems to involve". He comes to the conclusion that, within these terms, a way of life or a type of society can be judged as objectively better or worse than another. But his inadequate discussion of the complex concept of objectivity lets him down here. To claim that some moral rules (for example the prohibition of murder) are objectively valid in the sense that a rational, social life would be impossible in any society in which they were not accepted, is one thing; and we might draw the conclusion that we should be justified in enforcing these rules on recalcitrants and sceptics. But it is quite another thing to say that there is one type of life that is objectively better in the sense of being better for *everyone*. It was in this sense of ' better ' that Mill asked whether it is better to be a Socrates dissatisfied than a fool satisfied; but the right answer to his question is that, for many people. to emulate Socrates is to be a fool. The idea of a single good for man trapped Plato into thinking, not only (what may well be true) that only the just can be happy, but also (what is certainly false) that only philosophers can be. The doctrine of the good for man needs to be crossed with the very different doctrine that morality is concerned with an important, but essentially limited area of human life, the idea of a minimum set of indispensable rules for maintaining social order beyond which morality has nothing to say. I doubt whether Osborn would dissent from this liberal tradition; but his fear of relativism leads him to ignore it. We can surely admit that it is neither good nor wise for all of us to aim at the same pattern without "lapsing into subjectivism". Chapters 6-8, the

most technically philosophical parts of the book, though they contain some acute criticism of philosophical views from a psychological standpoint, are the least satisfactory. The discussion of meaning and the relation between questions about meaning and first-order ethical questions is not deep enough. The excursus into general epistemology in chapter 8 is weak : but this does not matter, since it is also unnecessary.

P. H. NOWELL-SMITH

Anaximander and the Origins of Greek Cosmology. By CHARLES H. KAHN. Columbia University Press, New York. Pp. xiii + 250. Price 52s.

DEBATE about the history of Presocratic thought is endless, for the fragmentary state of the evidence will always leave room for differences of opinion. Since such debate is highly interesting and enjoyable, this is no matter for regret, and Professor Kahn's thoughtful and well-documented study is a contribution to it of the first importance. It is built on the thesis that " from the sixth to the fourth century the Greek study of nature is a close-knit unity, and the originality of individual thinkers is intelligible only against the background of a continuous tradition, that is to say, of a common set of problems, principles, and solutions " (p. 207). The object is " to throw light on the essential unity of early Greek natural speculation " by emphasizing common traits rather than differences (p. 5).

That such a common basis existed is true, and Kahn has brilliantly succeeded in demonstrating it more clearly and making it more comprehensible. At the same time this deliberate throwing of the emphasis on one side occasionally tends to obscure differences and originality where they undoubtedly existed. Anaximander, as the writer of the first prose treatise on Nature (p. 240), is rightly made the leading figure. Yet throughout the book one is puzzled by the paucity of reference to what is generally regarded as his most original contribution, the idea that the *arche* or starting point of the universe was the boundless and undifferentiated, the *apeiron* as such. It finally comes in an appendix, where it is explained that the topic of *archai* or first principles has been largely neglected in the body of the book because " it conveniently serves to oppose the philosophers to one another ", whereas the author's concern has been to discover what they held in common. But if they were opposed on something so fundamental, does not this seriously weaken the thesis of " essential unity " on which the book takes its stand? Perhaps not, but at least this objection, which will be in the mind of any informed reader from the first, might have been met earlier.

Some may think also that the search for common features only has done some injustice to so outstandingly original a thinker as Heraclitus. On page 210 for instance we read: " Stoic cosmology is so intimately akin to that of Heraclitus (*and hence to that of the entire Ionian school*) . . ." (my italics), and page 151 perhaps gives a somewhat too " Heraclitean " view of the Ionian cosmologists. In this connection Kahn has revived the now unpopular opinion that in spite of fr. 30, Heraclitus believed the present cosmos to have a beginning and end in time (pp. 225 f.). This is part of a more general assertion, several times repeated, that no-one before Aristotle believed the world to be ungenerated. Yet to me it is clear from Aristotle's remarks in *Metaph.* A, 986 b 10-26, that in his eyes at least Xenophanes had believed this also, and I find nothing in the fragments

of Xenophanes to contradict the impression. As to Parmenides, it is of course only the Way of Seeming that is included in Kahn's generalization, though this is not explicitly pointed out. (See p. 225 n. 2.)

Another controversial point is the author's discounting of the originality of Pythagoras and treatment of the Italian school as " an offshoot of Ionian philosophy and not as its rival " (p. xi). The argument that "Anaximander is ancestor not only to the naturalism of Ionia, but also to the geometric philosophy usually associated with the name of Pythagoras " gives to certain geometrical aspects of his cosmology a prominence which is perhaps no more than their due : yet I cannot but feel that he was still a long way from the precise harmony of numbers which (with its exaltation of *peras*, not *apeiron*) was for the Pythagoreans the *arche* of the cosmos. Moreover this thesis ignores the whole religious side of the Western tradition (Empedocles as well as the Pythagoreans), which with its cycle of births, closely linked to the concept of the kinship of all nature, was an integral part of their system and foreign to the Ionian school.

To debate the many interesting points of detail would take too long. I will single out two only. On page 159 Alcmaeon's work is dated without comment at the end of the sixth century. This may be right, but since many scholars have placed him considerably later, some acknowledgment of their views is called for. On page 170 two passages from Herodotus are cited as instances of χρεών with the article, but in both cases τό is a relative pronoun.

This review has gone on the opposite principle from the book with which it is concerned : it has emphasized differences at the expense of a large measure of agreement and sympathy with the views expressed. To sum up, the work abounds in evidence of a keen historical sense and a deep insight into early ways of thinking which are all too rare among scholars who tackle this difficult and complex, yet infinitely important period of European thought about the natural world and its causes. This is perhaps most strikingly shown in the chapter on the origins and history of the doctrines of elements and opposites, from which every student of the Presocratics will find that he has much to learn.

W. K. C. GUTHRIE

A Modern Introduction to Moral Philosophy. By ALAN MONTEFIORE. Routledge. 14s.

MR. MONTEFIORE says in an introduction that his book is not intended for those who are already philosophers, professional or otherwise. This is surely too modest. Parts of the book—especially his discussion of the fact-value dichotomy and the nature of approval in Chapters 8 to 10—will certainly interest many such people.

The main subject—almost the only subject, preliminaries apart—is the nature of moral judgements. Mr. Montefiore opts for the familiar view that they are value judgements rather than statements or propositions, but that such judgements may nevertheless be called true or false.

Many people, however, will object to the rather untypical third chapter, in which Montefiore uses something like Hume's fork to suggest—he does not call it proof—that moral judgements are not statements.

The argument is basically that (p) all synthetic statements are *a posteriori*, while (q) moral judgements are neither *a posteriori* nor analytic.

Since (r) all statements are analytic or synthetic, it follows that moral judgements are not statements. If I understand, there are two reasons why Montefiore does not call this a proof, in spite of its formal validity. One is that we must take seriously the objective naturalist's contention that q is false in an important interpretation ; and Montefiore proceeds to do this in the four chapters already mentioned. The other is that p, q and r are merely analytic, merely guaranteed by a certain usage of the terms " analytic ", " synthetic ", " *a posteriori* ", and what Montefiore calls ' associated terms on either side of the analytic-synthetic fence '. The desirability of adopting this usage or terminology rather than one in which the contentions are non-analytic is said to be something which one cannot show in an introductory work, and the beginner is asked to be content with the reflection that many philosophically experienced people think one must adopt some such terminology for the sake of clarity.

This second idea, that p, q, and r are analytic in a terminology worthy of serious consideration, seems to be unconvincingly supported. In the first place, Montefiore does not even clearly mention—does not even show that there exists—a terminology which guarantees the truth of p, for instance. He claims, on page 33, that in the relevant terminology " . . . a proposition is counted as synthetic only if it excludes some logically possible state of affairs, and this is enough to render it *a posteriori*." Is it enough? Why is it enough?

Montefiore in effect defines " *a posteriori* " on page 28, as " such that its truth-value has to be ascertained by observation ". He also takes it for granted that a logically possible state of affairs is the same thing as a state of affairs describable without self-contradiction. Accepting those equivalences, the question then is whether there can be a statement which excludes a state of affairs describable without self-contradiction, but whose truth-value does not have to be ascertained by observation. Can there not be such a statement? Why can there not be?

We surely cannot construct a formal derivation of " *a posteriori* ", as defined, from " synthetic ", as defined. That being so, the question whether there can be such a statement must be settled by marshalling examples ; for instance, moral judgements. Unfortunately, a definite, supposedly convention-authenticated, negative answer to the question is being used by Montefiore to determine the status of those samples. His only qualification is that the convention might not be accepted ; but in fact, for all that we have seen, we must independently settle the status of moral judgements before we can say whether a negative answer *is* convention-authenticated.

Perhaps Montefiore means to close the gap between " synthetic " and " *a posteriori* " in a more arbitrary way. Perhaps he does not mean, as I have been taking him to mean, that " synthetic ", as defined by the group in question, clearly does entail " *a posteriori* ", as defined by that group ; in virtue of the ordinary meaning of the *definientes*. Perhaps he means that the group in question make it a basic rule to call nothing a ' statement that excludes a state of affairs that can be described without self-contradiction ' unless it is such that its truth-value has to be determined by observation.

If that is what he means, he is surely giving an unfeeling caricature of the group, whoever they are. For them, as for everyone else, the question, " does this exclude a state of affairs describable without contradiction? " is a question about a certain description, and about contradiction ; it asks, " Is contradiction involved in this description? " For them, as for everybody else, the question " Does the truth-value of this have to be ascertained

by observation? " is a question about observation, and whether it is necessary in order to ascertain the truth-value of a certain remark. That being so, the question whether all synthetic statements are *a posteriori* must, for them at least, be a question put to memory and imagination ; even though for some other group it might be a question answerable by truth-tables and definitional substitution.

Even if Montefiore had succeeded in mentioning a usage which confers analyticity on the claim that there are no synthetic *a priori*, would that not be a barren victory?

For one thing, it would have no tendency to show that he himself, or anyone else, actually uses the terms in accordance with that set of rules when engaged in philosophical thinking. (Some part in obscuring this obvious fact is played by the phrase " in *his* sense ". We tend to think that if someone says a certain doctrine is analytic when the crucial terms in it are used in his sense, then he cannot very well be wrong. It depends on what is meant by " in his sense ", however. Sometimes it may be clear that the doctrine is analytic for one who uses the terms in a sense which he has explained ; but it does not follow that he himself is such a one.)

Then again, if it could actually be shown that some group uses the relevant terms in such a way as to confer analyticity on the doctrine, would this not still be a trivial fact? Montefiore is prepared for this objection, though not, I think, for the logically prior objections just stated. (Unless I am quite mistaken, he falls into the not uncommon mistake of thinking that the philosophical claim to have secured some thesis against disproof by arbitrary definition is always clear, though of disputable value ; while in fact, in many contexts, such claims are simply unintelligible.) His reply to this last objection is that those who regard such a terminology as an indispensable instrument of clarification may well be right, since even those who believe there are synthetic *a priori* find it hard to say how there can be. Although this was not the sense of Kant's question as to how synthetic *a priori* are possible, there is some truth in the remark, if only because philosophy in this jargon of analyticity, etc., is very nearly ' the abuse of terms specially invented for the purpose '. Mr. Montefiore's reply here seems to accord ill, however, with his faith in the powers of arbitrary definition. If that faith were well placed, would not his synthetic a priorist opponent be entitled to reply as follows? " There is no difficulty. By 'analytic ' I mean the same as you do—knowable by analysis of the meaning alone. By ' *a posteriori* ' I also mean the same as you do—requiring to be tested by observation. We part company, however, in our usage of 'synthetic ' and ' *a priori* '. By the former I simply mean ' not analytic ', and by the latter, ' not *a posteriori* '. Consequently your arguments are arguments *for* the view that moral judgements are synthetic *a priori*, in *my* sense."

I have used my space to criticise what I think the weakest part of the book because a fundamental, and it seems still important, question of method is involved. Unfortunately this gives a rather misleading impression of the book. In the main it is a good book and a good introduction because in the other chapters a quite different, purely investigative, method is adopted.

One last remark about " true ", which leaves Mr. Montefiore with an unsolved puzzle as to whether it is an evaluative term in its application to statements and judgements. Surely the main point about truth is that after testing a soil-sample for acidity *or* deciding that an act was wrong we do not face the further task of testing for truth the statement that the soil-sample

was acid, *or* deciding whether the judgement, that the act was wrong, is true. In testing the soil-sample for acidity we *did* test for truth the statement that the soil-sample was acid, and in deciding that the act was wrong we *did* decide that the judgement that the act was wrong, was true. Consequently, there simply are no criteria for truth in the ordinary semantic sense of " criteria ". There are no true-making characteristics, either in the sense in which there are acid-making characteristics, or in the sense (assuming it to be a different one) in which there are wrong-making characteristics. If someone speaks of the " criteria " for truth he must be speaking, unsuitably, of the associations and pragmatics of this particular mode of assertion or assent ; but in that case it is hard to see where evaluation can come in.

<div style="text-align: right">J. M. HINTON</div>

Merit and Responsibility : A Study in Greek Values. By ARTHUR W. H. ADKINS. Clarendon Press : Oxford University Press. Pp. xvi + 380. 42s. net.

THAT there are radical differences between the presuppositions which underlie the moral philosophies of Plato and Aristotle, and those which are commonly accepted today, is a fact which should be obvious to any but the most uncritical reader of those authors. The precise nature of these differences, however, their causes, and the significant but frequently mystifying divergence in the choice and treatment of the problems to which they give rise, is by no means always so obvious to the philosophical and general reader, who may well approach fourth century Greek ethics with only a comparatively slight acquaintance with preceding and contemporary Greek literature. Dr. Adkins's book, written " in the hope that it might make clear to me why I could not understand the moral philosophy of Plato and Aristotle " (p. 1), should go a long way towards remedying this deficiency.

The scope of the work is indicated by the title of the doctoral thesis on which it is based : " The Development of the Concept of Moral Responsibility from Homer to Aristotle ". A distinction is drawn between two groups of values : the competitive, in terms of which a man is praised or blamed for his success or failure in maintaining or enhancing his own status, and that of his family or city, and the cooperative, which relate to those dealings between men, such as contracts, in which fairness is required. Since it is only in the latter context that the Greek held intentions to be relevant, while the highest term of praise possible, *agathos*, was almost invariably reserved for the man who excelled in the competitive virtues, it was inevitable that the concept of moral responsibility should play a much slighter role for the Greek than for us, with our predominantly Kantian morality.

The greater part of the book is an exposition, illustrated by a copious and carefully-arranged profusion of quotations, of the tensions arising within this system of values through the development of the city-state, with its need of the *agathos polites*, the declining belief (never very strongly held) in a divinely guaranteed moral order, and the influence of the Sophistic movement, with its subtle analyses of the concepts of *phusis* and *nomos*, and of cause. The endeavour to link possession of the major cooperative virtue, *dikaiosune*, with the *agathos*, reaches its culmination

in the writings of Plato; and the foregoing exposition clarifies one of the major puzzles of the *Republic*, that Plato attempts to prove, against the 'immoralist' Thrasymachus, that Justice is more *profitable* than Injustice —that it is in fact a prerequisite of success both for the individual and for the city. That Plato should conceive his problem in this way was inevitable, given the competiti connotations of *agathos*: but its dangerous implications are clear, particularly for the city-state, which is thereby encouraged to adopt any measures whatever which may be necessary to become just, and thus *agathos*, and to disregard the intentions of such of its citizens as obstruct it. In conclusion, Adkins examines the more loosely-knit moral philosophy of Aristotle, and its analysis of moral responsibility which is frequently regarded as a precursor of the contemporary approach to the subject; but argues that the overriding claim which *theoria* (which is not a cooperative virtue) is allowed to possess is equally inimical to the development of the concept in question along the lines which we should consider appropriate.

In the light of the largely negative conclusions of the book, as regards moral responsibility, it certainly might appear perverse (as Adkins himself admits) to have made that concept the central theme of a study of Greek values in general. While I agree with his rejoinder, that by focusing attention on a concept so important to us and so relatively unimportant to the Greek the significant differences between the two approaches to ethical problems are brought more clearly into view, I cannot but feel that the course of the discussion in some of the more finely-wrought passages would have been somewhat easier to follow had the focusing been done more often: perhaps a summary of the argument would have been a useful addition. In other respects, the organisation of the book is impressively clear, and the wealth of footnotes and the comprehensive indices will add to its usefulness for the specialist. It is to be hoped that the non-specialist will not be deterred by the frequent occurrence of Greek words (albeit transliterated) in the text from reading a work which should prove for a wide range of readers a valuable analysis of an interesting and important subject.

ANTHONY RALLS

Some Dilemmas of Naturalism. By W. R. DENNES. Columbia University Press (London: Oxford University Press.) Pp. 151. 32s.

THESE five Woodbridge Lectures (the sixth series) delivered in Columbia University are difficult to describe. The main structure of the argument is not easy to follow and the presentation is sometimes in the loose and easy-going style suitable to a lecture but sometimes so compact and allusive as to require slow and repeated reading.

It is not clear to me (though it may be to United States readers) what theories are included under 'naturalism'. The 'dilemmas' turn up sporadically throughout the book, and few of them receive any clear or specific solution. Some of them trigger off a discursive discussion but others are left in the air. Some, as Professor Dennes admits, are problems for any philosophy and not only for naturalism. Many of them cluster round the reflexive nature of philosophy. (1) Explaining the notions basic to interpretation requires the use of such notions. (2) The use of symbols can be discussed only by using symbols. (3) Any account of the

analytic-synthetic distinction has to rely on the analyticity it attempts to explicate. (4) Yet the analytic statements of philosophers then become merely stipulative definitions (*cf.* Ayer on ' meaning ' or Ryle and Austin on ' voluntary '). (5) Or alternatively they are empirical statements about how words or sentences are used, or value judgments about their ' best ' or ' proper ' use. If these dilemmas are regarded as insoluble we may fall back on the view that philosophic statements *tell* us nothing but only *show* how expressions are used.

The first three lectures discuss the issues raised by these questions. They present some definite and debated positions backed by very varying amounts of argument. (1) A verification theory of meaning commits its holder to a metaphysic of distinct events with qualities and relations. (2) If doubt is cast on such a metaphysic, it may be replied that (however questionable physical events and processes may be) I am aware of my own acts of intending and talking which are such qualitied and related processes. (3) The variation argument for the subjectivity of observed qualities can be met by maintaining that attention is being focused on different real factors. (4) An extensional theory of meaning cannot deal with open sets without reference to intention (the properties of members) which the speaker ' has in mind '. And ' having in mind ' is a private experience which may involve structured images. (5) Similarly, with action, intentions cannot be exhaustively described in terms of behaviour patterns ; there is an inner life of wants, images and self-discussions with reference to which action is directed. Besides these specific views there is a difficult, allusive and inconclusive section on the analytic-synthetic distinction.

All these points are interesting and deserve the fuller discussion which lecture limitations presumably precluded.

The last two lectures (on ' the Naturalistic Fallacy ' and ' Reason and Moral Imperatives ') contrast with the first three. Here the argument moves steadily forward in simple and easy terms. No new ground is broken, but a familiar current version of Hume's rejection of reason in morals is effectively presented. Factual premisses cannot lead to moral conclusions ; in any particular moral judgment the elements which admit of rational discussion are non-moral, the moral elements being expressions of subjective preference or approval. Dennes himself holds that moral choices harmonize the basic urges in man ; but, in case this looks like a naturalistic theory, he makes two points. First, this view of his is just his preference. Secondly, there are ideals as well as harmonious culture patterns, and sometimes, by reference to an ideal, harmony may be achieved by eliminating urges incompatible with the ideal, so that the urges are adapted to fit the pattern and not vice versa. Ideals themselves are also only expressions of preferences.

Dennes does not do much to illuminate two difficulties in approval theories such as his. How is moral approval to be distinguished from other types of approval? Because, says Dennes, it is more serious and involves not merely individual urges but a whole way of life ; it is not like choosing beef or mutton off a menu. But the choice of a career is serious and involves the preference of a whole way of life. Is this a moral choice? Secondly, Dennes does not make clear what exactly moral judgments express. He calls his last lecture ' Reason and Moral Imperatives ', yet he writes in terms of approval. He recognizes that approval fits best with the range of words allied to ' good ', but he has at

the end a page or two on 'right' and 'ought' and 'duty' (which fit best the imperative or prescriptive view). It would seem that his own solution would be a kind of compromise. Emphasis on 'goodness' leads to an open-minded awareness of the great range of values and the intricate variations of conduct if harmony is to be achieved. At its extreme this attitude rejects all rules and becomes an unlimited toleration, tempting one to inaction. Emphasis on 'rightness' indicates the influence of tradition resulting in conscientious adherence to ingrained rules and leading in the extreme case to rigidity and the view that in morals all is right and nothing good. But a compromise, Dennes admits, would again be only his particular preference.

J. D. MABBOTT

Religious Belief. By C. B. MARTIN. Cornell University Press, Ithaca (London : Oxford University Press), 1959. Pp. vii + 168. £1 4s.

Mr. MARTIN holds that when religious language, and the language of Christianity in particular, is taught to anyone " evasion and obscurity are present from the beginning. . . . There may be a temptation to think that religious statements at their rock bottom are in order. I do not believe this is the case " (p. 2). The disorderly character of religious language arises out of the reference it involves to God considered as something apart from the world. Theism can also (*cf.* Hume and Wisdom) be represented " mainly as a certain organization of facts about the world and not as a reference to some kind of ultimate Being outside of the world and our comprehension . . . " (p. 15). But this mode of interpretation is rejected by the author as a " pseudopantheistic reduction " inimical to the Christian tradition : it presents " an account of a substitute use so aseptic that it can make ' theists ' of us all " (p. 16).

Mr. Martin proceeds to diagnose incoherence in the idea that God is perfectly good, along with the idea that God is the source of all moral value and the idea that God is necessarily good. To argue that there is confusion in these theological ideas is not tantamount to arguing that there is confusion in Christianity : Christians are under no obligation to attach meaning to them, so far as I am aware, and the exposure of any nonsense there might be in them could do no possible disservice to theology. Mr. Martin's critique rests largely on the contention that the term " God " may be used as a proper name referring to a particular being, or else as a descriptive term, but not in both ways at once. An analysis is offered of the use of proper names which is illuminating in its own right.

Other areas in which Mr. Martin finds incoherence are the idea of encountering God in religious experience and the idea of surviving death. The first is considered mainly in relation to the opinions of John Baillie and H. H. Farmer, opinions which are not without interest, though I think it unduly flattering to these writers to suggest that their confusions are essential to religion. The discussion of life after death takes the customary line of focusing the difficulty upon personal identity ; also, objections are urged against the possibility of disembodied existence. In view of the fact that ' survival ' is being considered in the context of Christianity it is perhaps a pity that the resurrection of Christ should only receive a passing

mention. This is in a footnote with the remark : " It is not clear that we whose bodies shall be turned to dust are in the same position as Christ " (p. 107 n).

But if after being taken down from the Cross Christ's body had decomposed or turned to dust, the situation for the disciples would not have been materially different, except that the case for the resurrection might have become in one way easier by the preclusion of any possible lingering doubt as to whether Christ had died. Certainly, we are not in the same position as Christ, but that our bodies shall be turned to dust is beside the point here. Christians believe that they are to be resurrected. They are not committed to any particular theory of personal identity, *e.g.* that " there *must* be a something that is continuous and identical in this life that will survive into the next " (p. 107), an insistence argued convincingly by Mr. Martin to be pointless. Christians do not have a theory here at all. They believe that they are in for damnation or salvation, which is something Mr. Martin does not consider. The notion of looking " forward to life after death as a means of settling questions concerning the existence and nature of God " which he seems particularly concerned to attack (p. 95) smacks of ' Spiritualism ' perhaps rather than Christianity.

In a valuable chapter on Faith and Ritual, the author distinguishes between religious statements made *in the faith* and religious statements made *outside the faith*. He categorizes the former as ritualistic and compares them with ritualistic utterances outside religion. This chapter is followed by a pithy denunciation of the idea that assertion can exceed understanding. Mr. Martin pursues clarity with Cartesian passion and has no brief for " those who will complain, whenever anything is shown in a clear light, that seen clearly it is no longer what they saw in darkness " (p. vi). It behoves one who takes this line to watch out for himself, and I think that the question discussed by Mr. Martin in his last chapter—Why is there anything at all?—calls for a greater vigilance than he exercises. He treats theism as a hypothesis, although it is certainly no hypothesis as far as Christianity is concerned, and he regards " the difficulty of thinking clearly concerning any alternative cosmological theory " as " part of the strength of the theistic hypothesis " (p. 146). I think that this is loose ; that Mr. Martin should have written, not of the difficulty, but of the conceptual impossibility of framing alternative theories here, and that he should have taken this to argue confusion in the idea of theism as a hypothesis or cosmological theory. He says too that " we can conceive of what it would be like for nothing to exist. What we can understand is a possibility " (pp. 151-2, *cf.* p. 158 : " it seems to me that a possible state of affairs is that there should be nothing "). Go on then, tell us what it would be like ! I am not suggesting that it is impossible to do anything at all with the assertion that there might have been nothing : for instance, a person who said " There might have been nothing and then I should not be here " could be regarded as signifying the dependent character of his existence. But the assertion is empty until something *is* done with it ; and just contemplating it, entertaining the idea as a possibility, is not doing anything with it, or rather is trying to do something that cannot be done with it. You need a picture frame to contemplate the absence of a picture.

R. F. HOLLAND

Critique of Religion and Philosophy. By WALTER KAUFMANN. London, Faber and Faber, 1959. Pp. xvii + 325. 25s.

KAUFMANN has given us a very penetrating and critical, but not unsympathetic, analysis of religious thought and practice. His discussion, which is lively, extensive in scope and erudite, contains much to stimulate the philosopher and still more that will give the sincere Christian believer many headaches. In the earlier sections of his book Kaufmann considers the present-day philosophical scene, and he sees it as a scene of two revolts against the " abstractness and artificiality " which has often been evident in traditional philosophy. The two revolts are Existentialism and Analytic philosophy (Moore, Wittgenstein and followers)—movements which, different as they are, try to bring philosophy down to earth again. Kaufmann finds much to commend in both revolts, but complains that Existentialism is in danger of degenerating into bad poetry while Analytic philosophy is often content with being merely therapeutic. For him philosophy at its best means insight, vision and liberation, and of course he is not the first to emphasize the point (*cf.* Waismann's contribution to *Contemporary British Philosophy*, 3rd ser. ed. H. D. Lewis).

The bulk of this book concerns religion. The discussion, which ranges over such topics as truth (considered in connection with Zen Buddhism, Judaism and Christianity), revelation, miracles, causes of belief, proofs of God's existence, the doctrine of analogy and the significance of religious symbols, is far from inconsiderate but certainly does contain a forceful array of arguments which are critical of many traditional religious claims. The author is more openly belligerent in his attacks on Liberalism, Higher Criticism, and certain theologians such as Reinhold Niebuhr and Rudolph Bultmann. Some of the criticisms have already been made by the theologians themselves but Kaufmann brings them together in a discussion that is fresh and challenging. What he says is often incisive, sometimes caustic. Indeed some antagonistic bias is evident in his treatment of New Testament Christianity, and in particular one could justifiably quarrel with his interpretation of New Testament morality as prudential.

In considering what is the core of religion Kaufmann turns his attention to mysticism. William James's very inadequate criteria of mysticism are replaced by others which are far more plausible, but Kaufmann concludes that mysticism is not by any means the essence of religion. Religion is something more like loyalty to a tradition though it is not quite that either, nor is it simply morality. Basically it is poetry, but differs from other poetry in that it has become authoritarian in the form of dogma and ritual. The Bible is canonized poetry. Kaufmann thinks that Biblical scholars are misunderstanding Biblical literature completely when they try to rid it of inconsistencies and ambiguities, the theologians are wide off the mark when they try to systematize religious thought along rational lines, and religious believers generally are at fault in objectifying the divine.

We are left wondering what Kaufmann means by poetry, and it does not help to be told for instance that *philosophy* is poetry become self-critical, poetry with an intellectual conscience, or poetry subjected to rational scrutiny (p. 257). This does not throw much light on Kaufmann's view of philosophy or religion. In spite of this, what Kaufmann says in the course of his attempt to get at the core of religion is extremely suggestive and shows rare sensitivity and insight. That is why the book deserves to be read and re-read—especially by those who may want to reject half, or more, of what is said in it.

O. R. JONES

Received also :—

Aristotle De Anima ed. Sir David Ross, Oxford Clarendon Press (Oxford University Press), 1961, pp. 338, £2 10s.

R. Aron, *The Dawn of Universal History*, London, Weidenfeld & Nicholson, 1961, pp. 70, 7s. 6d.

E. Bettoni, *Duns Scotus : the basic principles of his philosophy*, trans. and ed. B. Bonansea, Washington, D.C., Catholic University of America Press, 1961, pp. 220, $4.75.

G. Boas, *The Limits of Reason*, London, Routledge & Kegan Paul, 1961, pp. 162, 12s. 6d.

I. M. Bocheński, *Contemporary European Philosophy*, University of California Press (Cambridge University Press), 1961, pp. xviii + 326, 15s. 6d.

R. B. Brandt, *Value and Obligation*, New York, Harcourt, Brace & World Inc., 1961, pp. x + 715, $6.95.

R. S. Brumbaugh, *Plato on the One*, The Hypotheses in the *Parmenides*, Yale University Press, 1961, pp. xiii + 364, $6.50.

J. Brun, *Aristote et la Lycée*, Paris, Presses Universitaires de France, 1961, pp. 128.

E. de Bruyne, *Grondproblemen van de Wijsgerige Logica*, Brussells, Paleis der Academiën, 1959, 2 vols., pp. 889.

G. Buchdahl, *The Image of Newton and Locke in the Age of Reason*, London, Sheed & Ward, 1961, pp. 116, 5s.

E. Cassirer, *The Logic of the Humanities*, trans. C. S. Howe, Yale University Press, pp. 217, $4.50.

E. Cassirer, *The Myth of State*, Yale University Press (London : Oxford University Press), 1961, pp. xii + 303, 11s. 6d.

Chambers, *Perception, Understanding and Society*, London, Sidgwick & Jackson, 1961, pp. 300, £1 5s.

Q. Gibson, *Facing Philosophical Problems*, Melbourne, F. W. Cheshire, 1961, pp. 63, 12s. 6d.

N. E. Christensen, *On the Nature of Meanings*, Copenhagen, Munksgaard, 1961, pp. 223, D.Kr. 28.

I. M. Copi, *Introduction to Logic*, second ed., London, Macmillan Co., 1961, pp. xviii + 512, £1 10s.

J. P. Day, *Inductive Probability*, London, Routledge & Kegan Paul, 1961, pp. xvi + 336, £1 5s.

Descartes Discours de la Méthode, ed. G. Gadoffre, Manchester University Press, 1961, pp. xliii + 117, 7s. 6d.

C. J. Ducasse, *The Belief in a Life after Death*, Springfield, Ill., C. C. Thomas (U.K. Basil Blackwell), 1961, pp. xvii + 318, £3 10s.

E. L. Fackenheim, *Metaphysics and Historicity* (Aquinas Lecture), Marquette University Press, 1961, pp. 120, $2.50.

E. J. Furlong, *Imagination*, London, George Allen & Unwin Ltd. (New York : The Macmillan Co.), 1961, pp. 125, 18s.

E. Gilson, *The Christian Philosophy of Saint Augustine*, London, Victor Gollancz Ltd., 1961, pp. xii + 398, £2 2s.

A. C. Graham, *The Problem of Value*, London, Hutchinson University Library, 1961, pp. 192, 12s. 6d.

J. Grenier, *Absolu et Choix*, Paris, Presses Universitaires de France, 1961, pp. 119, 4.50 N.F.

D. W. Hamlyn, *Sensation and Perception*, London, Routledge & Kegan Paul 1961, pp. xi + 210, £1 5s.

R. Harré, *Theories and Things*, London, Sheed & Ward, 1961, pp. 113, 5s.

Hegel, Studies in (Tulane Studies in Philosophy IX), Tulane University, 1960, pp. 187, $2.00.

I. L. Horowitz, *Radicalism and the Revolt against Reason*, The Social Theories of Georges Sorel, London, Routledge & Kegan Paul, 1961, pp. viii + 264, £1 10s.

J. Hospers, *Human Conduct*, An Introduction to the Problem of Ethics, New York, Harcourt, Brace & World Inc., 1961, pp. viii + 600, $6.50.

The Hsiao Ching trans. M. L. Makra, ed. P.K.T. Sih, New York, St. John's University Press, 1961, pp. xiv + 67, $3.50.

D. Joravsky, *Soviet Marxism, and the Natural Science 1917-1932*, London, Routledge & Kegan Paul Ltd., 1961, pp. 433, £2 5s.

R. Kroner, *Speculation and Revelation in Modern Philosophy*, Philadelphia, Westminster Press, 1961, pp. 316, $6.50.

R. Kuppi, *Über die Leibnizsche Logik*, Helsinki Societas Philosophica, 1960, pp. 279.

A. Lamouche, *La Théorie Harmonique Esthetique*, Paris, Dunod, 1961, pp. 356.

Lao Tzu/Tao Teh Ching trans. J. C. H. Wu, ed. P. K. T. Sih, New York, St. John's University Press, 1961, pp. xiv + 114.

A. Maslow, *Wittgenstein's Tractatus*, University of California Press, 1961, pp. xxii + 162, $4.00.

P. Merlan, *From Platonism to Neoplatonism*, second edition revised, The Hague, M. Nikhof, 1961, pp. xix + 250, 18 guilders.

H. Meyer, *Das Weltspiel*, Bern, Verlag Hans Huber, 1961, pp. 143, D.M. 11.50.

E. C. Moore, *American Pragmatism : Peirce, James and Dewey*, Columbia University Press (London : Oxford University Press), 1961, pp. xii + 285, £2.

R. Mukerjee, *The Philosophy of Social Science*, London, Macmillan & Co. Ltd., 1961, pp. ix + 187, £1 2s. 6d.

S. Nordentoft, *Heideggers Opgør med den Filosofiske tradition kritisk belyst*, Copenhagen, Hans Reitzel, 1961, pp. 187, D.Cr. 20,00.

S. C. Pepper, *World Hypotheses*, University of California Press (London : Cambridge University Press), 1961, pp. xiii + 348, 17s.

Philosophy of Science, Current Issues in, ed. H. Feigl and G. Maxwell, New York, Holt, Rinehart & Winston, 1961, pp. ix + 484.

J. J. Poortman, *De Grondparadox*, Assen, Van Gorcum, 1961, pp. 366, Hfl. 17.50.

J. Pucelle, *La Nature et l'esprit dans la Philosophie de T. H. Green*, Louvain, Éditions Nauwelaerts, 1961, pp. 327, F.B. 250.

K. W. Rankin, *Choice and Chance*, Oxford, Basil Blackwell, 1961, pp. 182, 18s.

Realism and the Background of Phenomenology, edited by R. M. Chisholm, London, George Allen & Unwin Ltd., 1961, pp. 308, £2 2s.

A. Roth, *Edmund Husserls Ethische Untersuchungen*, The Hague, Martinus Nijhoff, 1960, pp. xvii + 171.

P. F. Schmidt, *Religious Knowledge*, Glencoe, Ill., The Free Press, 1961, pp. ix + 147, $4.00.

D. W. Sherburne, *A Whiteheadian Aesthetic*, Yale University Press, 1961, pp. x + 219, $5.00.

E. Sillem, *Ways of Thinking about God*, London, Darton, Longman & Todd Ltd., 1961, pp. viii + 190, 11s. 6d.

M. G. Singer, *Generalisation in Ethics*, New York, A. A. Knopf, 1961, pp. xxvii + 351.

G. Smith and L. H. Kendzierski, *The Philosophy of Being*, Metaphysics I, New York, Macmillan Co., 1961, pp. xvii + 408, $ 5.75.

F. Solmsen, *Aristotle's System of the Physical World*, A Comparison with his predecessors (Cornell Studies in Classical Philology XXXIII), Cornell University Press (London : Oxford University Press), 1961, pp. xiv + 468, £3.

U. Spirito, *Inizio di una Nuova Epoca*, Firenze, G. C. Sansoni, 1961, pp. 340, L. 2.500.

S. Stebbing, *A Modern Elementary Logic*, London, Methuen, University Paperback, 1961, pp. x + 214, 7s. 6d.

A. E. Taylor, *Elements of Metaphysics*, London, Methuen, University Paperback, 1961, pp. 419, 12s. 6d.

R. K. Tripathi, *Spinoza in the light of the Vedanta*, Banaras Hindu University, 1961, pp. viii + 349, Rs. 12/-.

R. Tschumi, *A Philosophy of Literature*, London, Linden Press, 1961, pp. 190, £1 5s.

R. Tschumi, *Science, Philosophy and Literary Criticism*, Zürich, Polygraphischer Verlag, 1961, pp. 36.

C. A. Viano, *Scritti editi e inediti sulla Tolleranza*, Turin, Taylor, 1961, pp. 255, L. 2500.

Voltaire Philosophical Letters translated by E. Dilworth, New York, Liberal Arts Press, 1961, pp. xvi + 150, $1.00.

P. di Vona, *Studi sull'Ontologia di Spinoza*, Firenze, La Nuova Italia Editrice, 1961, pp. vii + 275, L. 2000.

J. Witt-Hansen, *Historical Materialism* Book I : Method, Copenhagen, Munksgaard, 1961, pp. 154, D.Kr. 38.

J. A. C. Brown, *Freud and the post-Freudians*, Pelican, 1961, pp. viii + 225, 3s. 6d.

Experimental Psychiatry, Lectures ed. by H. W. Brosin, University of Pittsburgh Press, 1961, pp. vi + 361, $7.50.

Sigmund Freud (1873-1939) Letters of, ed. E. L. Freud, London, Hogarth Press, 1961, pp. 464, £2 10s.

Fundamentals of Psychology : The Psychology of Thinking, Symposium of the New York Academy of Sciences, 1960, pp. 158, $3.00.

Handbook of Research Methods in Child Development, edited by P. H. Mussen, London, John Wiley & Sons Inc., 1961, pp. x + 1061, £6 2s.

Les Attitudes, Symposium by H. C. J. Duijker, P. Fraisse, R. Meili, P. Oléron, J. Paillard and others, Paris, Presses Universitaires de France, 1961, pp. 190, N.F. 10.

H. C. Lindgren and D. Byrne, *Psychology*. An Introduction to the Study of Human Behaviour, London, John Wiley & Sons Inc., 1961, pp. xi + 429.

The Manipulation of Human Behaviour, edited by A. D. Biderman and H. Zimmer, London, John Wiley & Sons Inc., 1961, pp. xii + 323, £3 4s.

Memory, Learning and Language, The Physical Basis of Mind, ed. W. Feindel, University of Toronto Press (London : Oxford University Press), 1961, pp. ix + 69, 16s.

O. H. Mowrer, *The Crisis in Psychiatry and Religion*, London, D. van Nostrand Co. Inc., 1961, pp. vii + 264, 5s.

J. Piaget, *Les Mécanismes Perceptifs*, Paris, Presses Universitaires de France, 1961, pp. 454, N.F. 28.

D. F. Pocock, *Social Anthropology*, London, Sheed & Ward, 1961, pp. 118, 5s.

The Psychoanalytic Study of the Child, Vol. XV, Hogarth Press, 1961, pp. 481, £2 10s.

Rorschach Psychology, ed. M. A. Rickers-Ovsiankina, London & New York, John Wiley & Sons, 1960, pp. xvi + 483, £3 8s.

S. Schachter, *The Psychology of Affiliation*, London, Tavistock Publications, 1961, pp. 141, £1 10s.

H. C. Shands, *Thinking and Psychotherapy*, Harvard University Press (London : Oxford University Press), 1961, pp. xiv + 319, £2 6s.

W. U. and B. J. Snyder, *The Psychotherapy of Relationship*, The Macmillan Co., 1961, pp. xi + 418, £2 12s. 6d.

A. Balakian, *Surrealism*, New York, Noonday Press, 1961, pp. 205, $1.45.

Robert Boyle, A Bibliography, by John F. Fulton, second edition, Oxford, Clarendon Press, 1961, pp. xvii + 217, £2 10s.

E. F. Caldin, *The Structure of Chemistry*, London, Sheed & Ward, 1961, pp. 49. 3s. 6d.

Canon Carpenter, *Common Sense about Christian Ethics*, London, Victor Gollancz Ltd., 1951, pp. 174, 12s. 6d.

L. K. Clark, *Pioneers of Prehistory in England*, London, Sheed & Ward, 1961, pp. 112, 5s.

D. Cox, *History and Myth*, London, Darton, Longman & Todd, 1961, pp. xiv + 166, 18s.

Cross, R., *Precedent in English Law* (Clarendon Law Series), Clarendon Press : Oxford University Press, 1961, pp. viii + 268, £1 1s.

N. W. Gilbert, *Renaissance Concepts of Method*, Columbia University Press (London : Oxford University Press), 1961, pp. xxvi + 255, £2 8s.

E. E. Hirschmann, *On Human Unity*, London, Victor Gollancz Ltd., 1961, pp. 237, £1 5s.

Il Problema della Demitizzazione, by various authors, ed. E. Castelli, Padua, CEDAM, 1961, pp. 334, L. 3.500.

R. Phukan (ed.), *The Sāmkhya Kārikā of Iśvarakrsna*, Calcutta, Firma K. L. Mukhopadhyay, 1961, pp. 165, 9s.

R. Powell, *Zen and Reality*, London, George Allen & Unwin Ltd., 1961, pp. 140, £1 1s.

A. Rapoport, *Fights, Games and Debates*, University of Michigan Press (London : Cresset Press Ltd.), 1961, pp. xvi + 400 £2 5s.

D. D. Runes, *Letters to my Teacher*, New York, Philosophical Library, 1961, pp. 106, $2.75.

F. Schuon, *Stations of Wisdom*, London, John Murray, 1961, pp. 157, £1 1s.

G. L. S. Shackle, *Decision Order and Time in Human Affairs*, Cambridge University Press, 1961, pp. xiv + 302, £1 15s.

Rabindranath Tagore, Pioneer in Education, exchanges between him and L. K. Elmhirst, London, John Murray, 1961, pp. 110, 15s.

R. K. Ullmann, *Tolerance and the Intolerable* Swarthmore Lecture 1961, London, George Allen & Unwin Ltd., 1961, pp. 74, 5s.

The Writer's Dilemma, by various authors, introduction by Stephen Spender, Oxford University Press, 1961, pp. xix + 88, 12s. 6d.

XI.—NOTES

MIND ASSOCIATION

Those who wish to join the Association should communicate with the Hon. Treasurer, Mr. J. D. MABBOTT, St. John's College, Oxford, to whom the yearly subscription of Twenty-five Shillings (payable in advance) should be sent. Cheques should be made payable to the Mind Association. Members may pay a Life Composition of £27 instead of the annual subscription. The annual subscription may be paid by Banker's Order; forms for this purpose can be obtained from the Hon. Treasurer.

The Hon. Secretary of the Association is Mr. A. R. WHITE, The University, Hull.

Members resident in U.S.A. are requested to pay the subscription ($4.00) to the Hon. Assistant-Treasurer, Prof.Alan Anderson, 993 Yale Station New Haven, Conn.

Those who join the Association after previously buying MIND direct from the publishers or through a bookseller are asked to inform the Hon. Treasurer of this (and of the name of the bookseller) when they join.

Professor Brand Blanshard has resigned his position as United States Assistant Treasurer of the Association, and has been succeeded by his colleague at Yale, Mr. Alan Anderson. When Professor Blanshard took on the work of collecting the Association's dollar subscriptions twenty years ago the Association had 30 United States members in its total of 300 members. His aim then was to raise the sales in the U.S. until U.S. members equalled those in the United Kingdom. It is a tribute to his energy that he has this year achieved his objective. The Association has 505 U.S. members as against 489 from Great Britain and Ireland in a total membership of 1145.

SPECIAL NOTICE TO CONTRIBUTORS

Contributors of Articles and Discussion Notes should in future always enclose a stamped and addressed envelope. Those from countries outside the United Kingdom should send international stamp coupons to cover return postage.

BACK NUMBERS OF *MIND*

The Librarian of Brasenose College, Oxford has for disposal the following back numbers of MIND : 1894 July, 1902 October, 1905 January, 1914 October, 1916 October, 1917 January and April, 1904 complete.

th the
whom
vance)
iation.
al sub-
Order;

g, The

ription
3 Yale

direct
rm the
y join.

States
by his
d took
y years
of 300
il U.S.
to his
ion has
a total

always
outside
o cover

ollowing
ry, 1914